



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 145331

TO: Janet Epps-Ford
Location: rem/2c05/2c18
Art Unit: 1635
Tuesday, March 01, 2005

Case Serial Number: 10/086062

From: Barb O'Brien
Location: Biotech-Chem Library
Remsen 1A69
Phone: 571-272-2518
barbara.obrien @uspto.gov

Search Notes

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2005, 23:32:36 ; Search time 729,221 Seconds
(without alignments)
1661.198 Million cell updates/sec

Title: US-10-086-062-1

Perfect score: 25
Sequence: 1 ctgagaccctctcgagagttccgct 25

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hig:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_srb:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	25	6	AX358106
2	25	100.0	813	6	AR097833
3	25	100.0	994	6	AR097835
4	25	100.0	1807	6	AR097836
5	25	100.0	2182	8	AY342393
6	25	100.0	2620	6	AR097837
7	25	100.0	2807	6	AX134224
8	25	100.0	2949	6	AX259248
9	25	100.0	3069	6	AX202417
10	25	100.0	3121	6	AX202416
11	25	100.0	3433	6	AR097838
12	25	100.0	3840	6	I06936
13	25	100.0	3840	6	I38571
14	25	100.0	3840	6	AR365568
15	25	100.0	3841	8	S94464
16	25	100.0	3972	6	AX355352
17	25	100.0	4190	8	ZMU29159
18	25	100.0	4341	6	AX259249
19	25	100.0	5465	6	BD194459

20	25	100.0	5522	12	AY452753
21	25	100.0	5622	6	AX039929
22	25	100.0	6103	6	BD194458
23	25	100.0	6111	6	BD247866
24	25	100.0	6111	6	AR193435
25	25	100.0	6111	6	AR430220
26	25	100.0	6215	6	AX118823
27	25	100.0	6217	6	AR209726
28	25	100.0	6579	6	AX118824
29	25	100.0	6667	6	AX118825
30	25	100.0	6818	6	AX840287
31	25	100.0	6898	6	BD132717
32	25	100.0	6898	6	AR209020
33	25	100.0	6898	6	AR482089
34	25	100.0	7018	12	AY452736
35	25	100.0	7027	6	AX840286
36	25	100.0	9172	6	AX259246
37	25	100.0	9359	6	AX384394
38	25	100.0	9359	6	AX473364
39	25	100.0	11643	6	AX840289
40	25	100.0	12817	6	AX356864
41	25	100.0	12849	6	BD247868
42	25	100.0	12949	6	AR193437
43	25	100.0	12949	6	AR430222
44	25	100.0	12982	6	BD091582
45	25	100.0	12982	6	BD091582

ALIGNMENTS

RESULT 1	AX358106	25 bp	DNA	linear	PAT 13-FEB-2002
LOCUS	AX358106	Sequence 1	from Patent WO0194394.		
DEFINITION	AX358106				
ACCESSION	AX358106.1	GI:18674853			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM	Agrobacterium tumefaciens (Rhizobium radiobacter)				
REFERENCE	Agrobacterium tumefaciens				
AUTHORS	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;				
TITLE	Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.				
JOURNAL	1 Jilka, J.M., Hood, E.B. and Howard, J.A.				
FEATURES	Novel plant promoter sequences and methods of use for same				
source	Patent: WO 0194394-A 1 13-DEC-2001;				
	Prodigene, Inc. (US)				
	Location/Qualifiers				
	1..25				
	/organism="Agrobacterium tumefaciens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:358"				
ORIGIN					
Query Match	100.0%; Score 25; DB 6; Length 25;				
Best Local Similarity	100.0%; Pred. No. 0.12; Mismatches 0; Indels 0; Gaps 0;				
Matches	25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1 CTGAGCCCTCTCGAGAGTTCGCT 25				
Db	1 CTGAGCCCTCTCGAGAGTTCGCT 25				
RESULT 2	AR097833	813 bp	DNA	linear	PAT 14-FEB-2001
LOCUS	AR097833	Sequence 13	from patent US 6072050.		
DEFINITION	AR097833				
ACCESSION	AR097833.1	GI:12806563			
VERSION					
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
	Unclassified.				

REFERENCE 1 (bases 1 to 813)
AUTHORS Bowen,B.A., Bruce,W.B., Lu,G., Sims,L.E. and Tagliani,L.A.
TITLE Synthetic promoters
JOURNAL Patent: US 6072050-A 13 06-JUN-2000;
FEATURES Location/Qualifiers
source 1..813
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 25; DB 6; Length 813;
Best Local Similarity 100.0%; Pred.No. 0.11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGACCCCTCTCGAGAGTTCGGCT 25
|||||
Db 654 CTGACCCCTCTCGAGAGTTCGGCT 678

RESULT 3
LOCUS AR097835 994 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 15 from patent US 6072050.
ACCESSION AR097835
VERSION AR097835.1 GI:12806565
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 994)
AUTHORS Bowen,B.A., Bruce,W.B., Lu,G., Sims,L.E. and Tagliani,L.A.
TITLE Synthetic promoters
JOURNAL Patent: US 6072050-A 15 06-JUN-2000;
FEATURES Location/Qualifiers
source 1..994
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 25; DB 6; Length 994;
Best Local Similarity 100.0%; Pred.No. 0.11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGACCCCTCTCGAGAGTTCGGCT 25
|||||
Db 654 CTGACCCCTCTCGAGAGTTCGGCT 678

RESULT 4
LOCUS AR097836 1807 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 16 from patent US 6072050.
ACCESSION AR097836
VERSION AR097836.1 GI:12806566
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1807)
AUTHORS Bowen,B.A., Bruce,W.B., Lu,G., Sims,L.E. and Tagliani,L.A.
TITLE Synthetic promoters
JOURNAL Patent: US 6072050-A 16 06-JUN-2000;
FEATURES Location/Qualifiers
source 1..1807
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 25; DB 6; Length 1807;
Best Local Similarity 100.0%; Pred.No. 0.11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGACCCCTCTCGAGAGTTCGGCT 25

Db 654 CTGACCCCTCTCGAGAGTTCGGCT 678

RESULT 5
LOCUS AY342393 2182 bp DNA linear PLN 01-AUG-2004
DEFINITION Zea diploperennis polyubiquitin-1 (Ubi-1) gene, promoter region and 5' UTR.
ACCESSION AY342393
VERSION AY342393.1 GI:37912415
KEYWORDS
SOURCE Zea diploperennis
ORGANISM Zea diploperennis
REFERENCE 1 (bases 1 to 2182)
AUTHORS Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
TITLE Analysis of the maize polyubiquitin-1 promoter heat shock elements and generation of promoter variants with modified expression characteristics
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2182)
AUTHORS Streetfield,S.J. and Love,R.T.
TITLE Direct Submission
JOURNAL Submitted (15-JUL-2003) Molecular Biology, ProdiGene, 101 Gateway Boulevard, College Station, TX 77845, USA
FEATURES Location/Qualifiers
source 1..2182
/organism="Zea diploperennis"
/mol_type="genomic DNA"
/db_xref="taxon:4576"
gene 1..>2182
/gene="Ubi-1"
promoter 1..922
/gene="Ubi-1"
mRNA 923..>1003
/gene="Ubi-1"
/product="polyubiquitin-1"
5'UTR 923..1003
/gene="Ubi-1"

ORIGIN

Query Match 100.0%; Score 25; DB 8; Length 2182;
Best Local Similarity 100.0%; Pred.No. 0.11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGACCCCTCTCGAGAGTTCGGCT 25
|||||
Db 708 CTGACCCCTCTCGAGAGTTCGGCT 732

RESULT 6
LOCUS AR097837 2620 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 17 from patent US 6072050.
ACCESSION AR097837
VERSION AR097837.1 GI:12806567
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2620)
AUTHORS Bowen,B.A., Bruce,W.B., Lu,G., Sims,L.E. and Tagliani,L.A.
TITLE Synthetic promoters
JOURNAL Patent: US 6072050-A 17 06-JUN-2000;
FEATURES Location/Qualifiers
source 1..2620
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 25; DB 6; Length 2620;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTGACCCCTCTCGAGAGTTCGGCT 25
DB 654 CTGACCCCTCTCGAGAGTTCGGCT 678

RESULT 7
AX134224 2807 bp DNA linear PAT 29-MAY-2001
LOCUS AX134224
DEFINITION Sequence 1 from Patent WO0132897.
ACCESSION AX134224
VERSION AX134224.1 GI:14270967
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..2807
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="synthetic construct"

ORIGIN
Query Match 100.0%; Score 25; DB 6; Length 2807;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTGACCCCTCTCGAGAGTTCGGCT 25
DB 1501 CTGACCCCTCTCGAGAGTTCGGCT 1525

RESULT 8
AX259248 2949 bp DNA linear PAT 26-OCT-2001
LOCUS AX259248
DEFINITION Sequence 46 from Patent WO0173087.
ACCESSION AX259248
VERSION AX259248.1 GI:16508494
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..2949
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="artificial sequence ZmUbi-GFP-35S term expression cassette"

ORIGIN
Query Match 100.0%; Score 25; DB 6; Length 2949;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTGACCCCTCTCGAGAGTTCGGCT 25
DB 692 CTGACCCCTCTCGAGAGTTCGGCT 716

RESULT 9
AX202417 3069 bp DNA linear PAT 30-AUG-2001
LOCUS AX202417
DEFINITION Sequence 5 from Patent WO0152620.
ACCESSION AX202417
VERSION AX202417.1 GI:15392164
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..3069
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="PND3018 coding region"

ORIGIN
Query Match 100.0%; Score 25; DB 6; Length 3069;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTGACCCCTCTCGAGAGTTCGGCT 25
DB 677 CTGACCCCTCTCGAGAGTTCGGCT 701

RESULT 10
AX202416 3121 bp DNA linear PAT 30-AUG-2001
LOCUS AX202416
DEFINITION Sequence 4 from Patent WO0152620.
ACCESSION AX202416
VERSION AX202416.1 GI:15392163
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..3121
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="PND3008 coding region"

ORIGIN
Query Match 100.0%; Score 25; DB 6; Length 3121;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTGACCCCTCTCGAGAGTTCGGCT 25
DB 677 CTGACCCCTCTCGAGAGTTCGGCT 701

RESULT 11
AR097838 3433 bp DNA linear PAT 14-FEB-2001
LOCUS AR097838
DEFINITION Sequence 18 from patent US 6072050.
ACCESSION AR097838

VERSION AR097838.1 GI:12806568
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 3433)
TITLE Bowen,B.A., Bruce,W.B., Lu,G., Sims,L.E. and Tagliani,L.A.
JOURNAL Synthetic promoters
Patent: US 6072050-A 18 06-JUN-2000;
FEATURES Location/Qualifiers
source 1..3433
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 25; DB 6; Length 3433;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGAGCCCTCTCGAGAGTTCGGCT 25
Db 654 CTGAGCCCTCTCGAGAGTTCGGCT 678

RESULT 12
LOCUS 106936 3840 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 8 from Patent EP 0342926.
ACCESSION 106936
VERSION 106936.1 GI:589816
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 3840)
Quail,P.H., Christensen,A.H., Hershey,H.P., Sharrock,R.A. and Sullivan,T.D.
JOURNAL Plant ubiquitin promoter system
Patent: EP 0342926-A2 8 23-NOV-1989;
FEATURES Location/Qualifiers
source 1..3840
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 25; DB 6; Length 3840;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGAGCCCTCTCGAGAGTTCGGCT 25
Db 686 CTGAGCCCTCTCGAGAGTTCGGCT 710

RESULT 13
LOCUS 138571 3840 bp DNA linear PAT 13-MAY-1997
DEFINITION Sequence 1 from patent US 5614399.
ACCESSION 138571
VERSION 138571.1 GI:2084625
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 3840)
Quail,P.H., Christensen,A.H., Hershey,H.P., Sharrock,R.A. and Sullivan,T.D.
JOURNAL Plant ubiquitin promoter system
Patent: US 5614399-A 1 25-MAR-1997;
FEATURES Location/Qualifiers
source 1..3840
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 25; DB 6; Length 3840;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGAGCCCTCTCGAGAGTTCGGCT 25
Db 686 CTGAGCCCTCTCGAGAGTTCGGCT 710

RESULT 14
LOCUS AR365568 3840 bp DNA linear PAT 03-SEP-2003
DEFINITION Sequence 1 from patent US 5510474.
ACCESSION AR365568
VERSION AR365568.1 GI:34429351
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 3840)
Quail,P.H., Christensen,A.H., Hershey,H.P., Sharrock,R.A. and Sullivan,T.D.
JOURNAL Plant ubiquitin promoter system
Patent: US 5510474-A 1 23-APR-1996;
FEATURES Location/Qualifiers
source 1..3840
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 25; DB 6; Length 3840;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGAGCCCTCTCGAGAGTTCGGCT 25
Db 686 CTGAGCCCTCTCGAGAGTTCGGCT 710

RESULT 15
LOCUS S94464 3841 bp DNA linear PLN 07-MAY-1993
DEFINITION polyubiquitin [maize, Genomic, 3841 nt].
ACCESSION S94464
VERSION S94464.1 GI:248336
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 3841)
Christensen,A.H., Sharrock,R.A. and Quail,P.H.
Maize polyubiquitin genes: structure, thermal perturbation of
expression and transcript splicing, and promoter activity following
transfer to protoplasts by electroporation
JOURNAL Plant Mol. Biol. 18 (4), 675-689 (1992)
MEDLINE 92216044
PUBMED 1313711
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI gisdbq 94464] from the original journal article.
FEATURES Location/Qualifiers
source 1..3841
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
1..3841
/gene="polyubiquitin"
1994..3595
/gene="polyubiquitin"
/codon_start=1

```

/protein_id="AAB21993.1"
/db_xref="GI:248337"
/translation="MQIFVKLTIGKTTITLEVESSDTIDNVKAKIQDKEGIPPDQORLI
FAGKQLEDGRTLADYNIQKESTLHLVLRGGMQIFVKLTIGKTTITLEVESSDTIDN
KAKIQDKEGIPPDQORLI FAGKQLEDGRTLADYNIQKESTLHLVLRGGMQIFVKLT
IKTITITLEVESSDTIDNVKAKIQDKEGIPPDQORLI FAGKQLEDGRTLADYNIQKEST
LHLVLRGGMQIFVKLTIGKTTITLEVESSDTIDNVKAKIQDKEGIPPDQORLI FAGK
QLEDGRTLADYNIQKESTLHLVLRGGMQIFVKLTIGKTTITLEVESSDTIDNVKAKI
QDKEGIPPDQORLI FAGKQLEDGRTLADYNIQKESTLHLVLRGGMQIFVKLTIGKT
ITLEVESSDTIDNVKAKIQDKEGIPPDQORLI FAGKQLEDGRTLADYNIQKESTLHLV
LRLRGGMQIFVKLTIGKTTITLEVESSDTIDNVKAKIQDKEGIPPDQORLI FAGKQLED
GRTLADYNIQKESTLHLVLRGGMQ"
    
```

ORIGIN

```

Query Match      100.0%; Score 25; DB 8; Length 384;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    
```

```

QY      1 CTGACCCCTCTCGAGAGTTCGCT 25
          |||||
Db      687 CTGACCCCTCTCGAGAGTTCGCT 711
    
```

Search completed: March 1, 2005, 00:29:10
 Job time : 735.221 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2005, 23:26:41 ; Search time 218.182 Seconds
(without alignments)
678.304 Million cell updates/sec

Title: US-10-086-062-1

Perfect score: 25
Sequence: 1 ctgacccctctcgaggttcgcgc 25

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	25	6	AAD24066 Maize wll
2	25	100.0	813	2	AAZ11619 Maize ubi
3	25	100.0	813	8	ACC78118 Maize ubi
4	25	100.0	994	2	AAZ11621 Ubi1 upst
5	25	100.0	1413	8	ACC78116 Nematode-
6	25	100.0	1807	2	AAZ11622 UCP2 prom
7	25	100.0	2005	10	ADC84542 Modified
8	25	100.0	2100	9	ACF06306 5' genome
9	25	100.0	2609	8	ACC78115 Nematode-
10	25	100.0	2620	2	AAZ11623 Ubi1 upst
11	25	100.0	2807	4	AAH19412 SUC-1 pro
12	25	100.0	2848	9	ACF06318 Pat inser
13	25	100.0	3069	4	AAZ11579 PND3018 r
14	25	100.0	3121	2	AAZ11578 PND3008 a
15	25	100.0	3433	2	AAZ11578 UCP4 prom
16	25	100.0	3840	1	AAH92365 Sequence
17	25	100.0	3972	6	ABT07353 Ecdysone
18	25	100.0	3972	10	ADF49176 Ecdysone
19	25	100.0	4341	4	AAZ19839 ZmUbi-GIG
20	25	100.0	4569	4	AAZ19838 ZmUbi-GFP

21	25	100.0	4587	9	ACF06316 Intact vi
22	25	100.0	5465	2	AAZ20088 Plasmid p
23	25	100.0	5622	3	AAZ06506 DNA conat
24	25	100.0	6103	2	AAZ20087 Plasmid p
25	25	100.0	6111	3	AAH54211 Transform
26	25	100.0	6215	4	AAZ03876 NotI frag
27	25	100.0	6579	4	AAZ03877 NotI frag
28	25	100.0	6667	4	AAZ03878 NotI frag
29	25	100.0	6818	10	ACF58342 Nucleotid
30	25	100.0	6898	2	AAV63742 Plasmid U
31	25	100.0	7027	10	ACF58341 Nucleotid
32	25	100.0	8352	12	ADP73862 Plasmid p
33	25	100.0	9144	12	ADP73861 Plasmid p
34	25	100.0	9172	4	AAZ19836 Plasmid p
35	25	100.0	9359	6	ABK11039 PVDH36 v
36	25	100.0	9359	6	ABK10687 Transform
37	25	100.0	11643	10	ACF58344 Nucleotid
38	25	100.0	12380	12	ADP73863 Plasmid p
39	25	100.0	12817	6	ABK15668 Expressio
40	25	100.0	12949	3	AAH54213 Transform
41	25	100.0	12982	4	AAH86432 Vector ps
42	25	100.0	12982	4	AAH86432 Vector ps
43	25	100.0	13269	6	AAZ03961 Maize ubi
44	25	100.0	13274	6	ABK15666 Binary ve
45	25	100.0	13829	11	ADN08781 Maize opt

ALIGNMENTS

RESULT 1
AAD24066
ID AAD24066 standard; DNA; 25 BP.
XX
AC AAD24066;
XX
DT 09-APR-2002 (first entry)
XX
DB Maize wild-type Ubi-1 promoter heat shock element.
XX
KW Gene expression; maize; ubiquitin promoter; Ubi-1; HSE;
KW heat shock element; agronomic gene; ds.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT misc_feature 1..15
FT FT /*tag= a
FT FT /*note= "5' heat shock element"
FT FT /*tag= b
FT FT /*note= "3' heat shock element"
XX
PN WO200194394-A2.
XX
PD 13-DEC-2001.
XX
PF 08-JUN-2001; 2001WO-US018689.
XX
PR 09-JUN-2000; 2000US-00590558.
XX
PA (PROD-) PRODIGENE INC.
XX
PI Jilka JM, Hood BE, Howard JA;
XX
DR WPI; 2002-122117/16.
XX
PT New promoter sequences for causing expression of a structural gene
PT especially agronomic gene or open reading frame in a plant cell,
PT comprises engineered versions of the maize ubiquitin promoter.
XX
PS Example 1; Page 43; 68p; English.
XX

CC The invention relates to a promoter sequence capable of directing
CC expression of a nucleotide sequence in a plant cell, comprising maize
CC ubiquitin (Ubi-1) promoter sequence with a modification so that it does
CC not include two overlapping heat shock elements (HSE) or it directs
CC expression to increase the endosperm/embryo expression ratio of the
CC protein when compared to the ratio from a wild-type ubiquitin promoter.
CC The modified Ubi-1 promoter comprises a deletion of 3', 5' or both HSEs,
CC two non-overlapping/adjacent HSEs, replacement of HSEs with a trimer of a
CC seed specific element from the promoter of pea lectin gene pal, or
CC insertion of a transcription factor binding site in the HSE region. An
CC expression construct comprising modified Ubi-1 promoter is useful for
CC causing expression of a structural gene (agronomic genes) or open reading
CC frame in a plant cell. The modified Ubi-1 promoter increases expression
CC levels beyond those observed with native ubiquitin promoter. The present
CC sequence is maize wild-type Ubi-1 promoter heat shock element
XX

SQL Sequence 25 BP; 3 A; 10 C; 6 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGAGCCCTCTCGAGAGTCCGCT 25
DB 1 CTGAGCCCTCTCGAGAGTCCGCT 25

RESULT 2
AA211619
ID AA211619 standard; DNA; 813 BP.
XX
AC AA211619;
XX
DT 16-NOV-1999 (first entry)
XX

DE Maize Ubi1 upstream activating region (UAR).
XX
XX Plant promoter; TATA motif; transcription start site; upstream element;
KW gene expression; oxalate oxidase; plant resistance; pathogen; maize;
KW Ubi-1 promoter; upstream activating region; ss.
XX

OS Zea sp.
XX

PN MO9943838-A1.
XX

PD 02-SEP-1999.
XX

PF 23-FEB-1999; 99WO-US003863.
XX

PR 24-FEB-1998; 98US-00028819.
XX

PA (PION-) PIONEER HI-BRED INT INC.
XX

PI Bowen BA, Bruce WB, Lu G, Sims LE, Tagliani LA;
XX

DR WPI; 1999-540601/45.
XX

PT New synthetic promoter functional in plants to provide non-tissue
PT specific, constitutive expression, particularly of oxalate oxidase for
PT increased resistance to pathogens.
XX
XX Disclosure; Page 51; 61pp; English.

CC The invention provides a new synthetic plant promoter that comprises a
CC TATA motif; a transcription start site (TSS) and a region between TATA
CC and TSS containing at least 64 percent GC content. The synthetic core
CC promoter, optionally containing additional upstream elements are used to
CC increase expression, provides non-tissue specific, constitutive
CC transcription of heterologous genes in any sort of plant, especially the
CC gene for oxalate oxidase for increasing plant resistance to pathogens.
CC The upstream activating elements can be used to increase transcription
CC from any promoter. A combination of the synthetic core promoter with
CC synthetic upstream elements can induce expression 10 times greater than

CC that provided by the maize Ubi-1 promoter. The present sequence
CC represents a maize Ubi1 upstream activating region (UAR)
XX

SQL Sequence 813 BP; 241 A; 166 C; 142 G; 264 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 2; Length 813;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGAGCCCTCTCGAGAGTCCGCT 25
DB 654 CTGAGCCCTCTCGAGAGTCCGCT 678

RESULT 3
ACC78118
ID ACC78118 standard; DNA; 813 BP.
XX
AC ACC78118;
XX
DT 18-AUG-2003 (first entry)
XX

DE Maize ubiquitin-1 promoter upstream activating region (UAR).
XX

KW Nematode; SCP1; UCP3; SUP; nematocidal; gene therapy; promoter;
KW ubiquitin-1; upstream activating region; UAR; maize; ds.
XX

OS Zea mays.
XX

PN WO2003033651-A2.
XX

PD 24-APR-2003.
XX

PF 08-OCT-2002; 2002WO-US032059.
XX

PR 16-OCT-2001; 2001US-0329667P.
XX

PA (PION-) PIONEER HI-BRED INT INC.
XX

PI Hu X, Lu G, Ruff RL, Schuh W;
XX

DR WPI; 2003-393512/37.
XX

PT New nucleic acid molecule comprising a heterologous nematode-resistance
PT sequence operatively linked to a promoter capable of driving
PT transcriptions of the sequence, useful for promoting nematode-resistance
PT in plants.
XX
XX Claim 16; Page 55; 65pp; English.

CC The invention relates to a nucleic acid molecule comprising a
CC heterologous nematode-resistance sequence operably linked to a promoter
CC capable of driving transcriptions of the heterologous nematode-resistance
CC sequence in a plant cell, where the promoters are selected from synthetic
CC SCP1, UCP3 and SUP promoters. The nucleic acid molecules and methods are
CC useful for promoting nematode-resistance in plants by inducing expression
CC of a heterologous nematode-resistance sequence, altering expression of a
CC heterologous nucleotide sequence, or creating or enhancing disease
CC resistance in a plant. The present sequence represents the upstream
CC activating region of (UAR) of the ubiquitin-1 promoter from maize
XX

SQL Sequence 813 BP; 241 A; 166 C; 142 G; 264 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 8; Length 813;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGAGCCCTCTCTCGAGAGTCCGCT 25
DB 654 CTGAGCCCTCTCTCGAGAGTCCGCT 678

RESULT 4

```

AA211621
ID AA211621 standard; DNA; 994 BP.
XX
XX AA211621;
AC
XX
XX 16-NOV-1999 (first entry)
DT
XX
XX Ubi1 upstream activating region (UAR) linked with a promoter fragment.
DE
XX
XX Plant promoter; TATA motif; transcription start site; upstream element;
KM gene expression; oxalate oxidase; plant resistance; pathogen; maize;
KM Ubi-1 promoter; upstream activating region; ss.
XX
XX Synthetic.
OS
XX Zea sp.
XX
XX WO9943838-A1.
XX
XX 02-SEP-1999.
PD
XX
XX 23-FEB-1999; 99WO-US003863.
XX
XX 24-FEB-1998; 98US-00028819.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Bowen BA, Bruce WB, Lu G, Sims LE, Tagliani LA;
XX
XX WPI; 1999-540601/45.
XX
XX New synthetic promoter functional in plants to provide non-tissue
PT specific, constitutive expression, particularly of oxalate oxidase for
PT increased resistance to pathogens.
XX
XX Disclosure; Page 52-53; 61pp; English.
XX
XX The invention provides a new synthetic plant promoter that comprises a
XX TATA motif; a transcription start site (TSS) and a region between TATA
XX and TSS containing at least 64 percent GC content. The synthetic core
XX promoter, optionally containing additional upstream elements are used to
XX increase expression, provides non-tissue specific, constitutive
XX transcription of heterologous genes in any sort of plant, especially the
XX gene for oxalate oxidase for increasing plant resistance to pathogens.
XX The upstream activating elements can be used to increase transcription
XX from any promoter. A combination of the synthetic core promoter with
XX synthetic upstream elements can induce expression 10 times greater than
XX that provided by the maize Ubi-1 promoter. The present sequence
XX represents a copy of Ubi1 upstream activating region (UAR) operably
XX linked with the core promoter
XX
XX Sequence 994 BP; 288 A; 209 C; 181 G; 316 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 25; DB 2; Length 994;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTGAGACCCCTCTCGAGAGTTCCGCT 25
DB 654 CTGAGACCCCTCTCGAGAGTTCCGCT 678

```

```

XX
XX Synthetic.
OS
XX Cauliflower mosaic virus.
OS
XX Zea mays.
XX
XX WO2003033651-A2.
XX
XX 24-APR-2003.
PD
XX
XX 08-OCT-2002; 2002WO-US032059.
XX
XX 16-OCT-2001; 2001US-0329667P.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Hu X, Lu G, Ruff RL, Schuh W;
XX
XX WPI; 2003-393512/37.
XX
XX New nucleic acid molecule comprising a heterologous nematode-resistance
PT sequence operatively linked to a promoter capable of driving
PT transcriptions of the sequence, useful for promoting nematode-resistance
PT in plants.
XX
XX Claim 1; Page 54-55; 65pp; English.
XX
XX The invention relates to a nucleic acid molecule comprising a
XX heterologous nematode-resistance sequence operably linked to a promoter
XX capable of driving transcriptions of the heterologous nematode-resistance
XX sequence in a plant cell, where the promoters are selected from synthetic
XX SCPI, UCP1 and SUP promoters. The nucleic acid molecules and methods are
XX useful for promoting nematode-resistance in plants by inducing expression
XX of a heterologous nematode-resistance sequence, altering expression of a
XX heterologous nucleotide sequence, or creating or enhancing disease
XX resistance in a plant. The present sequence represents a nematode-
XX regulated promoter SUP1, a synthetic hybrid promoter consisting of one
XX copy of the upstream activating region of (UAR) of the 35S CamV promoter
XX from Cauliflower mosaic virus operably linked to the Rayn7-Syn II core element which
XX is operably linked to one copy of the UAR of ubiquitin-1 promoter from Z.
XX mays, which is operably linked to the Rayn7-Syn II core promoter
XX
XX Sequence 1413 BP; 419 A; 309 C; 280 G; 405 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 25; DB 8; Length 1413;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTGAGACCCCTCTCGAGAGTTCCGCT 25
DB 1086 CTGAGACCCCTCTCGAGAGTTCCGCT 1110

```

```

RESULT 5
ACCT8116
ID ACCT8116 standard; DNA; 1413 BP.
XX
XX ACCT8116;
AC
XX
XX 18-AUG-2003 (first entry)
DT
XX
XX Nematode-regulated synthetic promoter SUP.
DE
XX
XX Nematode; SCPI, UCP3, SUP; nematocidal; gene therapy; promoter;
KM 35S promoter; ubiquitin-1; Rayn7-Syn II core promoter; ss.
XX

```

```

RESULT 6
AA211622
ID AA211622 standard; DNA; 1807 BP.
XX
XX AA211622;
AC
XX
XX 16-NOV-1999 (first entry)
DT
XX
XX UCP2 promoter sequence.
DE
XX
XX Plant promoter; TATA motif; transcription start site; upstream element;
KM gene expression; oxalate oxidase; plant resistance; pathogen; maize;
KM Ubi-1 promoter; upstream activating region; ss.
XX
XX Synthetic.
OS
XX Zea sp.
XX
XX WO9943838-A1.
XX
XX 02-SEP-1999.
XX

```

PF 23-FEB-1999; 99WO-US003863.
XX
XX 24-FEB-1998; 98US-00028819.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Bowen BA, Bruce WB, Lu G, Sims LE, Tagliani LA;
XX WPI; 1999-540601/45.
DR
XX
PT New synthetic promoter functional in plants to provide non-tissue
PT specific, constitutive expression, particularly of oxalate oxidase for
PT increased resistance to pathogens.
XX
XX
PS Disclosure; Page 53-54; 61pp; English.
XX
XX The invention provides a new synthetic plant promoter that comprises a
CC TATA motif; a transcription start site (TSS) and a region between TATA
CC and TSS containing at least 64 percent GC content. The synthetic core
CC promoter, optionally containing additional upstream elements are used to
CC increase expression, provides non-tissue specific, constitutive
CC transcription of heterologous genes in any sort of plant, especially the
CC gene for oxalate oxidase for increasing plant resistance to pathogens.
CC The upstream activating elements can be used to increase transcription
CC from any promoter. A combination of the synthetic core promoter with
CC synthetic upstream elements can induce expression 10 times greater than
CC that provided by the maize Ubi-1 promoter. The present sequence
CC represents a UCP2 promoter sequence (2 copies of Ubi1 upstream activating
CC region (UAR) operably linked with the core promoter)
XX
SQ Sequence 1807 BP; 530 A; 375 C; 323 G; 579 T; 0 U; 0 Other;
Query Match 100.0%; Score 25; DB 2; Length 1807;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTGAGACCCCTCTCGAGAGTTCCGCT 25
DB 654 CTGAGACCCCTCTCGAGAGTTCCGCT 678
RESULT 7
AD8C4542
ID AD8C4542 standard; DNA; 2005 BP.
XX
XX AC AD8C4542;
XX
XX 01-JAN-2004 (first entry)
XX
XX Modified maize ubiquitin promoter encoding sequence.
XX
XX expression cassette; ds.
XX
XX Zea mays.
XX
XX WO2003027257-A2.
XX
XX 03-APR-2003.
XX
XX 27-SEP-2002; 2002WO-US030895.
XX
XX 27-SEP-2001; 2001US-0325607P.
XX
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
XX Levin JZ, Dietrich R, Budziszewski GJ;
XX WPI; 2003-354651/33.
XX
XX New isolated nucleic acid molecule encoding a polypeptide with a 3'-5'
PT exonuclease domain, useful in molecular biology and transformation, in
PT particular for reproducing and predictably manipulating gene expression
PT in a plant cell.

XX
XX Disclosure; Page 105-106; 108pp; English.
PS
XX
XX The present invention relates to an invention that results in the
CC expression of the endogenous nucleotide sequence in plant cell being
CC increased. In this method the plant cell comprises a second expression
CC cassette comprising a nucleic acid molecule, where the expression of the
CC nucleic acid molecule of interest in the plant cell is decreased as
CC compared to expression of the nucleic acid molecule of interest in a
CC plant cell lacking the first expression cassette. The methods and
CC compositions of the present invention are useful in molecular biology and
CC transformation, in particular for reproducing and predictably
CC manipulating gene expression in a plant cell. The present sequence is a
CC nucleotide sequence of modified maize ubiquitin promoter.
XX
XX
SQ Sequence 2005 BP; 474 A; 426 C; 412 G; 693 T; 0 U; 0 Other;
Query Match 100.0%; Score 25; DB 10; Length 2005;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTGAGACCCCTCTCGAGAGTTCCGCT 25
DB 690 CTGAGACCCCTCTCGAGAGTTCCGCT 714
RESULT 8
ACF06306
ID ACF06306 standard; DNA; 2100 BP.
XX
XX ACF06306;
XX
XX 07-OCT-2003 (first entry)
XX
XX 5' genome+intact vip3a insert sequence SEQ ID NO:6.
XX
XX Maize; corn: VIP1034; vip3a; pat; vegetative insecticidal protein; VIP;
XX transgenic genotype; insect resistance; herbicide tolerance; corn event;
XX plant; gene; ds.
XX
XX Zea mays.
XX
XX WO2003052073-A2.
XX
XX 26-JUN-2003.
XX
XX 16-DEC-2002; 2002WO-US040099.
XX
XX 17-DEC-2001; 2001US-0341456P.
XX
XX 08-JAN-2002; 2002US-0346660P.
XX
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
XX Steiner H, Dawson J, Dunder E, Meghji M;
XX WPI; 2003-533014/50.
XX
XX New transgenic corn event, VIP1034, useful for conferring a
PT characteristic consisting of insect resistance and herbicide tolerance.
PT
XX Claim 7; Page 56-57; 73pp; English.
PS
XX
XX The present invention describes an isolated nucleic acid molecule (1)
CC comprising a sequence having at least one junction sequence of corn event
CC VIP1034 comprising a sequence having 20 bp, or its complement. VIP1034 is
CC a vegetative insecticidal protein (VIP), and comprises a novel transgenic
CC genotype that comprises a vip3a gene and a pat gene, which confers insect
CC resistance and herbicide tolerance to the transgenic corn event. Also
CC described: (1) an amplicon comprising (i); (2) a polynucleotide primer
CC for detecting corn event VIP1034 in a sample comprising at least 15
CC contiguous nucleotides from position 1-716 of the 2100-bp sequence given
CC in ACF06306 or its complement; (3) detecting the presence of a DNA
CC corresponding to the VIP1034 event in the sample; (4) a kit for detecting

CC the presence of VIP1034 nucleic acids in a sample; (5) detecting corn
CC event VIP1034 in a biological sample; (6) a seed of any corn inbred
CC comprising the transgenic genotype of the corn event VIP1034, where the
CC transgenic genotype comprises an intact copy and a fragmented copy of a
CC first expression cassette and an intact copy of a second expression
CC cassette; (7) a corn plant or its pollen or ovule comprising the
CC transgenic genotype of the corn event VIP1034; (8) a hybrid corn seed
CC produced by crossing the corn plant with an inbred corn plant having a
CC different genotype; (9) a hybrid corn plant produced by growing the
CC hybrid corn seed; (10) producing corn seed; and (11) producing hybrid
CC corn seeds. The nucleic acid is useful for conferring a characteristic
CC consisting of insect resistance and herbicide tolerance. The present
CC sequence represents a corn 5' genome-intact vip3A insert sequence, which
CC is used in the exemplification of the present invention

SQ Sequence 2100 BP; 555 A; 479 C; 397 G; 668 T; 0 U; 1 Other;

Query Match 100.0%; Score 25; DB 9; Length 2100;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGACCCCTCTCGAGAGTTCGCT 25
DB 1401 CTGACCCCTCTCGAGAGTTCGCT 1425

RESULT 9
ACC78115
ID ACC78115 standard; DNA; 2609 BP.

AC ACC78115;
XX
DT 18-AUG-2003 (first entry)

DE Nematode-regulated synthetic promoter UCP3.
XX
XX Nematode; SCP1; UCP3; SUP; nematocidal; gene therapy; promoter;
KM ubiquitin-1; Rayn7-Syn II core promoter; ss.
XX
XX Synthetic.
OS Zea mays.
XX
XX WO2003033651-A2.
PN 24-APR-2003.
PD
XX
XX 08-OCT-2002; 2002WO-US032059.
PF
XX
XX 16-OCT-2001; 2001US-0329667P.
PR
XX
XX (PION-) PIONEER HI-BRED INT' INC.
PA
XX
XX Hu X, Lu G, Ruff RL, Schuh W;
PI
XX
XX WPI; 2003-393512/37.
DR
XX
XX New nucleic acid molecule comprising a heterologous nematode-resistance
PT sequence operatively linked to a promoter capable of driving
PT transcriptions of the sequence, useful for promoting nematode-resistance
PT in plants.
XX
XX
XX Claim 1; Page 53-54; 65pp; English.

CC The invention relates to a nucleic acid molecule comprising a
CC heterologous nematode-resistance sequence operably linked to a promoter
CC capable of driving transcriptions of the heterologous nematode-resistance
CC sequence in a plant cell, where the promoters are selected from synthetic
CC SCP1, UCP3 and SUP promoters. The nucleic acid molecules and methods are
CC useful for promoting nematode-resistance in plants by inducing expression
CC of a heterologous nematode-resistance sequence, altering expression of a
CC heterologous nucleotide sequence, or creating or enhancing disease
CC resistance in a plant. The present sequence represents a nematode-
CC regulated promoter UCP3, a synthetic hybrid promoter consisting of three

CC copies (in forward orientation) of the upstream activating region (UAR)
CC of the ubiquitin-1 promoter from Z. mays operably linked to the Rayn7-Syn
CC II core promoter
XX
XX
SQ Sequence 2609 BP; 769 A; 542 C; 464 G; 834 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 8; Length 2609;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGACCCCTCTCGAGAGTTCGCT 25
DB 654 CTGACCCCTCTCGAGAGTTCGCT 678

RESULT 10
AAZ11623
ID AAZ11623 standard; DNA; 2620 BP.

AC AAZ11623;
XX
DT 16-NOV-1999 (first entry)

DE Ubi1 upstream activating region (UAR) linked with a promoter fragment.
XX
XX
XX Plant promoter; TATA motif; transcription start site; upstream element;
KM gene expression; oxalate oxidase; plant resistance; pathogen; maize;
KM Ubi-1 promoter; upstream activating region; ss.
XX
XX Synthetic.
OS Zea sp.
XX
XX WO9943838-A1.
PN 02-SEP-1999.
PD
XX
XX 23-FEB-1999; 99WO-US003863.
PF
XX
XX 24-FEB-1998; 98US-00028819.
PR
XX
XX (PION-) PIONEER HI-BRED INT' INC.
PA
XX
XX Bowen BA, Bruce WB, Lu G, Sims LB, Tagliani LA;
PI
XX
XX WPI; 1999-540601/45.
DR
XX
XX New synthetic promoter functional in plants to provide non-tissue
PT specific, constitutive expression, particularly of oxalate oxidase for
PT increased resistance to pathogens.
PT
XX
XX Disclosure; Page 54-56; 61pp; English.

CC The invention provides a new synthetic plant promoter that comprises a
CC TATA motif, a transcription start site (TSS) and a region between TATA
CC and TSS containing at least 64 percent GC content. The synthetic core
CC promoter, optionally containing additional upstream elements are used to
CC increase expression, provides non-tissue specific, constitutive
CC transcription of heterologous genes in any sort of plant, especially the
CC gene for oxalate oxidase for increasing plant resistance to pathogens.
CC The upstream activating elements can be used to increase transcription
CC from any promoter. A combination of the synthetic core promoter with
CC synthetic upstream elements can induce expression 10 times greater than
CC that provided by the maize Ubi-1 promoter. The present sequence
CC represents a copy of Ubi1 upstream activating region (UAR) operably
CC linked with the core promoter

SQ Sequence 2620 BP; 772 A; 541 C; 465 G; 842 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 2; Length 2620;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGACCCCTCTCGAGAGTTCGCT 25

DB 654 CTGACCCCTCTCGAGAGTTCCGCT 678

RESULT 11

AAH19412

AAH19412 standard; DNA; 2807 BP.

AC AAH19412;

30-JUL-2001 (first entry)

SUC-1 promoter element.

Promoter; SUC-1; transgenic plant; CamV 35S promoter; maize; polyubiquitin promoter; UBI promoter; sugarcane gene expression; monocotyledonous plant gene expression; ds.

OS Synthetic.

Key Location/Qualifiers

5'UTR 1715..1797

/note= "Maize UBI 5' UTR"

intron 1798..2807

/tag= c

WO200132897-A2.

10-MAY-2001.

06-NOV-2000; 2000WO-1B001601.

05-NOV-1999; 99ZA-00006954.

11-MAY-2000; 2000ZA-00002313.

(SASU-) SOUTH AFRICAN SUGAR ASSOC.

(UWST-) UNIV STELLENBOSCH INST PLANT BIOTECHNOLO.

Greenewald J, Botha FC;

WPI; 2001-316449/33.

A new promoter construct designated SUC-1 contains promoter elements from CamV 35S and maize UBI and is useful to confer stable high level constitutive expression in transgenic plants.

Claim 1; Fig 1; 19pp; English.

The present sequence is tandem promoter construct, SUC-1. This promoter confers stable, high level, constitutive expression in transgenic plants. The promoter contains both the cauliflower mosaic virus (CamV) 35S promoter and the maize polyubiquitin (UBI) promoter. The promoter is particularly useful for gene expression in monocotyledonous plants, such as sugarcane. The combination of the two promoters gives higher expression levels in monocotyledonous plants than the CamV 35S or UBI promoter when used alone

Sequence 2807 BP; 751 A; 607 C; 582 G; 867 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 4; Length 2807;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CTGACCCCTCTCGAGAGTTCCGCT 25

1501 CTGACCCCTCTCGAGAGTTCCGCT 1525

RESULT 12
ACF06318/c
ID ACF06318 standard; DNA; 2848 BP.
XX

AC ACF06318;

07-OCT-2003 (first entry)

Pat Insect nucleotide sequence SEQ ID NO:18.

Maize; corn; VIP1034; vip3a; pat; vegetative insecticidal protein; VIP; transgenic genotype; insect resistance; herbicide tolerance; corn event; plant; gene; ds.

Zea mays.

WO2003052073-A2.

26-JUN-2003.

16-DEC-2002; 2002WO-US040099.

17-DEC-2001; 2001US-0341456P.

08-JAN-2002; 2002US-0346660P.

(SYGN) SYNGENTA PARTICIPATIONS AG.

Steiner H, Dawson J, Dunder E, Meghji M;

WPI; 2003-533014/50.

New transgenic corn event, VIP1034, useful for conferring a characteristic consisting of insect resistance and herbicide tolerance. Claim 22; Page 68-69; 73pp; English.

The present invention describes an isolated nucleic acid molecule (I) comprising a sequence having at least one junction sequence of corn event VIP1034 comprising a sequence having 20 bp, or its complement. VIP1034 is a vegetative insecticidal protein (VIP), and comprises a novel transgenic genotype that comprises a vip3a gene and a pat gene, which confers insect resistance and herbicide tolerance to the transgenic corn event. Also described: (1) an amplicon comprising (1); (2) a polynucleotide primer for detecting corn event VIP1034 in a sample comprising at least 15 contiguous nucleotides from position 1-716 of the 2100-bp sequence given in ACF06306 or its complement; (3) detecting the presence of a DNA corresponding to the VIP1034 event in the sample; (4) a kit for detecting the presence of VIP1034 nucleic acids in a sample; (5) detecting corn event VIP1034 in a biological sample; (6) a seed of any corn inbred comprising the transgenic genotype of the corn event VIP1034, where the transgenic genotype comprises an intact copy and a fragmented copy of a first expression cassette and an intact copy of a second expression cassette; (7) a corn plant or its pollen or ovule comprising the transgenic genotype of the corn event VIP1034; (8) a hybrid corn seed produced by crossing the corn plant with an inbred corn plant having a different genotype; (9) a hybrid corn plant produced by growing the hybrid corn seed; (10) producing corn seed, and (11) producing hybrid corn seeds. The nucleic acid is useful for conferring a characteristic consisting of insect resistance and herbicide tolerance. The present sequence represents a corn pat insect nucleotide sequence, which is used in the exemplification of the present invention

Sequence 2848 BP; 937 A; 632 C; 580 G; 699 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 9; Length 2848;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CTGACCCCTCTCGAGAGTTCCGCT 25

2159 CTGACCCCTCTCGAGAGTTCCGCT 2135

RESULT 13
AAD11579
ID AAD11579 standard; DNA; 3069 BP.
XX

AC AAD11579;
 XX
 DT 24-SEP-2001 (first entry)
 XX
 DE pND3018 repression construct.
 XX
 KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
 KW modulation; plant technology; agriculture; pND3018 repression construct;
 KW ds.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT 44..2026
 FT /tag= a
 FT /label= Zmubi_promoter
 FT misc_feature 2066..2173
 FT /tag= b
 FT /note= "Sint interaction domain (SID) repression domain"
 FT protein_bind 2207..2735
 FT /tag= d
 FT /bound_molecy= "ZFP"
 FT sig_peptide 2620..2641
 FT /tag= c
 FT /note= "Nuclear localisation signal"
 FT misc_feature 2762..2791
 FT /tag= e
 FT /note= "HA epitope tag"
 FT terminator 2820..3112
 FT /tag= f
 FT /label= Nos_terminator
 XX
 PN WO200152620-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 19-JAN-2001; 2001WO-US001817.
 XX
 PR 21-JAN-2000; 2000US-0177468P.
 PR 21-JUL-2000; 2000US-00620897.
 XX
 PA (SCRI) SCRIPPS RES INST.
 PA (SYGN) SYNGENTA AGRIC DISCOVERY INC.
 XX
 PI Barbas CF, Stege JT, Guan X, Dalmia B;
 XX
 WI: 2001-465325/50.
 XX
 PT New zinc finger proteins, useful for modulating or regulating gene
 PT expression and metabolic pathways in plants, e.g. for treating in the
 PT plant cells a disorder that is associated with abnormal expression of the
 PT target gene.
 XX
 PS Claim 121; Page 134-136; 156pp; English.
 XX
 CC The parent discloses methods and compositions to modulate the expression
 CC of a target gene in plant cells. The method involves providing plant
 CC cells with a zinc finger protein (ZFP) which is capable of specifically
 CC binding to a target nucleotide sequence or its complementary strand
 CC within a target gene and allowing the ZFP binding to the target
 CC nucleotide sequence, where the expression of the target gene in the plant
 CC cells is modulated. The ZFP and fusions of the ZFP proteins are useful
 CC for modulating or regulating gene expression and metabolic pathways in
 CC plants. The ZFP, fusion proteins and methods are useful in plant and
 CC agricultural technology. The method is useful particularly for treating a
 CC disorder in the plant cells, where the disorder is associated with
 CC abnormal expression of the target gene. The present DNA sequence is
 CC pND3018 repression construct. This sequence is used to construct maize
 CC ubiquitin promoter Zmubi effector construct which is further used to
 CC determine the function of ZFP-effector fusion protein on luciferase
 CC reporter gene in maize cells
 CC
 CC Sequence 3069 BP; 747 A; 713 C; 669 G; 940 T; 0 U; 0 Other;
 SQ

Query Match 100.0%; Score 25; DB 4; Length 3069;
 Best Local Similarity 100.0%; Pred. No. 0.17;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 CTGACCCCTCTCGAGAGTTCGCT 25
 DB 677 CTGACCCCTCTCGAGAGTTCGCT 701
 RESULT 14
 AAD11578
 ID AAD11578 standard; DNA; 3121 BP.
 XX
 AC AAD11578;
 XX
 DT 24-SEP-2001 (first entry)
 XX
 DE pND3008 activation construct.
 XX
 KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
 KW modulation; plant technology; agriculture; pND3008 activation construct;
 KW ds.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT 44..2026
 FT /tag= a
 FT /label= Zmubi_promoter
 FT misc_feature 2060..2588
 FT /tag= b
 FT /note= "Six finger ZFP2C7"
 FT sig_peptide 2620..2641
 FT /tag= c
 FT /note= "Nuclear localisation signal"
 FT misc_feature 2641..2805
 FT /tag= d
 FT /note= "VP64 activation domain"
 FT misc_feature 2805..2836
 FT /tag= e
 FT /note= "HA epitope tag"
 FT terminator 2884..3164
 FT /tag= f
 FT /note= "Nos terminator"
 XX
 PN WO200152620-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 19-JAN-2001; 2001WO-US001817.
 XX
 PR 21-JAN-2000; 2000US-0177468P.
 PR 21-JUL-2000; 2000US-00620897.
 XX
 PA (SCRI) SCRIPPS RES INST.
 PA (SYGN) SYNGENTA AGRIC DISCOVERY INC.
 XX
 PI Barbas CF, Stege JT, Guan X, Dalmia B;
 XX
 WI: 2001-465325/50.
 XX
 PT New zinc finger proteins, useful for modulating or regulating gene
 PT expression and metabolic pathways in plants, e.g. for treating in the
 PT plant cells a disorder that is associated with abnormal expression of the
 PT target gene.
 XX
 PS Example 1B; Page 132-134; 156pp; English.
 XX
 CC The parent discloses methods and compositions to modulate the expression
 CC of a target gene in plant cells. The method involves providing plant
 CC cells with a zinc finger protein (ZFP) which is capable of specifically
 CC binding to a target nucleotide sequence or its complementary strand

CC within a target gene and allowing the ZFP binding to the target
 CC nucleotide sequence, where the expression of the target gene in the plant
 CC cells is modulated. The ZFP and fusions of the ZFP proteins are useful
 CC for modulating or regulating gene expression and metabolic pathways in
 CC plants. The ZFP, fusion proteins and methods are useful in plant and
 CC agricultural technology. The method is useful particularly for treating a
 CC abnormal expression of the target gene. The present DNA sequence is
 CC PND3008 activation construct. This sequence is used to construct maize
 CC ubiquitin promoter Zmubi effector construct which is further used to
 CC determine the function of ZFP-effector fusion protein on luciferase
 CC reporter gene in maize cells

SQ Sequence 3121 BP; 755 A; 713 C; 684 G; 969 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 4; Length 3121;

Best Local Similarity 100.0%; Pred. No. 0.17; Mismatches 0; Gaps 0;

Matches 25; Conservative 0; Indels 0; Gaps 0;
 QY 1 CTGAGCCCTCTCGAGAGTTCGCT 25
 DB 677 CTGAGCCCTCTCGAGAGTTCGCT 701

RESULT 15

AA211624 ID AA211624 standard; DNA; 3433 BP.

AA211624; AC

16-NOV-1999 (first entry) DT

UCP4 promoter sequence. DE

plant promoter; TATA motif; transcription start site; upstream element;
 gene expression; oxalate oxidase; plant resistance; pathogen; maize;
 Ubi-1 promoter; upstream activating region; ss.

Synthetic. OS

Zea sp. OS

MO9943838-A1. PN

02-SEP-1999. PD

23-FEB-1999; 99WO-US003863. PF

24-FEB-1998; 98US-00028819. PR

(PION-) PIONEER HI-BRED INT INC. PA

Bowen BA, Bruce WB, Lu G, Sims LE, Tagliani LA; PI

WPI; 1999-540601/45. DR

New synthetic promoter functional in plants to provide non-tissue
 specific, constitutive expression, particularly of oxalate oxidase for
 increased resistance to pathogens.

Disclosure, Page 56-58; 61pp; English. PS

CC The invention provides a new synthetic plant promoter that comprises a
 CC TATA motif; a transcription start site (TSS) and a region between TATA
 CC and TSS containing at least 64 percent GC content. The synthetic core
 CC promoter, optionally containing additional upstream elements are used to
 CC increase expression, provides non-tissue specific, constitutive
 CC transcription of heterologous genes in any sort of plant, especially the
 CC gene for oxalate oxidase for increasing plant resistance to pathogens.
 CC The upstream activating elements can be used to increase transcription
 CC from any promoter. A combination of the synthetic core promoter with
 CC synthetic upstream elements can induce expression 10 times greater than
 CC that provided by the maize Ubi-1 promoter. The present sequence
 CC represents a UCP4 promoter sequence (4 copies of Ubi1 upstream activating

CC region (UAR) operably linked with the core promoter)

SQ Sequence 3433 BP; 1014 A; 707 C; 607 G; 1105 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 2; Length 3433;
 Best Local Similarity 100.0%; Pred. No. 0.17; Mismatches 0; Gaps 0;

Matches 25; Conservative 0; Indels 0; Gaps 0;
 QY 1 CTGAGCCCTCTCGAGAGTTCGCT 25
 DB 654 CTGAGCCCTCTCGAGAGTTCGCT 678

Search completed: February 28, 2005, 23:51:37
 Job time : 223.182 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: February 28, 2005, 23:35:01 / Search time 1734.42 Seconds
(without alignment)
548.661 Million cell updates/sec

Title: US-10-086-062-1

Perfect score: 25
Sequence: 1 ctgacccctctcagagatccgc 25

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

EST:
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	229	8	BH638797
2	25	100.0	307	9	CG805015
3	25	100.0	338	9	CG805036
4	25	100.0	534	8	BH638999
5	25	100.0	620	9	CG709678
6	25	100.0	620	9	CG805054
7	25	100.0	716	7	CG472745
8	25	100.0	765	7	CG472714
9	25	100.0	769	7	CG472736
10	25	100.0	774	7	CG472678
11	25	100.0	805	7	CG472728
12	25	100.0	817	7	CG472729
13	25	100.0	819	7	CG472694
14	25	100.0	820	7	CG472674
15	25	100.0	825	7	CG472692
16	25	100.0	828	7	CG472722
17	25	100.0	828	7	CG472731
18	25	100.0	833	7	CG472737
19	25	100.0	837	7	CG472693
20	25	100.0	841	7	CG472673
21	25	100.0	851	7	CG472751
22	25	100.0	852	7	CG472735
23	25	100.0	852	7	CG472741
24	25	100.0	856	7	CG472725

25	25	100.0	859	7	CG472744	CG472744	GQ0081.BR
26	25	100.0	876	7	CG472681	CG472681	GQ0081.BR
27	25	100.0	879	7	CG472711	CG472711	GQ0081.BR
28	25	100.0	940	9	CG670569	CG670569	CGUFM88TV
29	25	100.0	965	9	CG342774	CG342774	CGYB89TV
30	23	92.0	794	7	CG472764	CG472764	GQ0081.BR
31	23	92.0	853	7	CG472754	CG472754	GQ0081.BR
32	21	84.0	720	8	CK438248	CK438248	GQ0081.BR
33	20.2	80.8	100	8	BH219170	BH219170	1006085A0
34	20.2	80.8	415	4	BM61080	BM61080	952044B06
35	20.2	80.8	553	4	BM736979	BM736979	952051F11
36	20.2	80.8	636	4	BE79714	BE79714	949078B10
37	20.2	80.8	728	9	CG212795	CG212795	CGXCY79TV
38	20.2	80.8	896	9	CG212788	CG212788	CGXCY79TV
39	19.8	79.2	655	6	C55594	C55594	YUJ1
40	19.2	76.8	669	9	CL173701	CL173701	1043771
41	19.2	76.8	897	5	BU838185	BU838185	AGENCOURT
42	19.2	76.8	908	2	BE797311	BE797311	601587675
43	19.2	76.8	1283	8	AZ130054	AZ130054	OSJNB010
44	18.8	75.2	323	4	BI029471	BI029471	IL5-MT026
45	18.8	75.2	472	2	BE142500	BE142500	RC6-HT014

ALIGNMENTS

RESULT 1
LOCUS BH638797 229 bp DNA linear GSS 14-FEB-2002
DEFINITION 1008024H11.2EL_x1 1008 - Rescuemu Grid I Zea mays genomic, genomic survey sequence.
ACCESSION BH638797
VERSION BH638797.1 GI:18663263
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 229)

REFERENCE
AUTHORS Walbot, V.
TITLE Maize genomic sequences found using engineered Rescuemu transposon
JOURNAL Unpublished (2001)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu

Possible ligation site of ends cut by 2 different endonucleases.
Reverse complemented post-ligation sequence from source sequence.
Plate: 1008024 row: 10
Classes: transposon-tagged.
Location/Qualifiers
1..229

FEATURES

source
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_11b="1008 - Rescuemu Grid I"
/note="Organ: leaf; Vector: Rescuemu (engineered from pBluescript backbone); Site_1: BamHI; Site_2: BglII; Rescuemu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on Rescuemu, go to the web site www.zmdb.iastate.edu and follow the links for 'Rescuemu.' Grid I was grown at Berkeley in 2001. DNA was extracted from leaf punches, double digested using BamHI

and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

ORIGIN

Query Match 100.0%; Score 25; DB 8; Length 229;
Best Local Similarity 100.0%; Pred. No. 0.64;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CTGAGACCCCTCTCGAGAGTTCCGCT 25
DB 25 CTGAGACCCCTCTCGAGAGTTCCGCT 49

RESULT 2
LOCUS CG805015 307 bp DNA linear GSS 10-NOV-2003
DEFINITION 118056F03.2EL_x1 1118 - Rescuedu Grid S Zea mays genomic, genomic survey sequence.
ACCESSION CG805015 GI:38241016
VERSION CG805015.1
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD 1 (bases 1 to 307)
Walbot, V.
Maize genomic sequences found using engineered Rescuedu transposon
REFERENCE 1 (bases 1 to 307)
AUTHORS Walbot, V.
TITLE Maize genomic sequences found using engineered Rescuedu transposon
JOURNAL Unpublished (2001)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Possible ligation site of ends cut by 2 different endonucleases.
Reverse complemented post-ligation sequence from source sequence.
Plate: 118056 row: 40
Class: transposon-tagged.
Location/Qualifiers
1..307
/organism="Zea mays"
/mol_type="genomic DNA"
/culturvar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1118 - Rescuedu Grid S"
/notes="Organ: leaf; Vector: Rescuedu (engineered from pBluescript backbone); Site 1: BamHI; Site 2: BglII; Rescuedu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on Rescuedu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'Rescuedu'. Grid S was grown at San Diego in 2002. DNA was extracted from leaf strips, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

FEATURES

source

ORIGIN

Query Match 100.0%; Score 25; DB 9; Length 307;
Best Local Similarity 100.0%; Pred. No. 0.65;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CTGAGACCCCTCTCGAGAGTTCCGCT 25
DB 80 CTGAGACCCCTCTCGAGAGTTCCGCT 104

RESULT 3
LOCUS CG805036 338 bp DNA linear GSS 10-NOV-2003
DEFINITION 118056F03.2EL_x1 1118 - Rescuedu Grid S Zea mays genomic, genomic survey sequence.
ACCESSION CG805036
VERSION CG805036.1 GI:38241053
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD 1 (bases 1 to 338)
Walbot, V.
Maize genomic sequences found using engineered Rescuedu transposon
REFERENCE 1 (bases 1 to 338)
AUTHORS Walbot, V.
TITLE Maize genomic sequences found using engineered Rescuedu transposon
JOURNAL Unpublished (2001)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Possible ligation site of ends cut by 2 different endonucleases.
Reverse complemented post-ligation sequence from source sequence.
Plate: 118056 row: 40
Class: transposon-tagged.
Location/Qualifiers
1..338
/organism="Zea mays"
/mol_type="genomic DNA"
/culturvar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1118 - Rescuedu Grid S"
/notes="Organ: leaf; Vector: Rescuedu (engineered from pBluescript backbone); Site 1: BamHI; Site 2: BglII; Rescuedu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on Rescuedu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'Rescuedu'. Grid S was grown at San Diego in 2002. DNA was extracted from leaf strips, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

FEATURES

source

ORIGIN

Query Match 100.0%; Score 25; DB 9; Length 338;
Best Local Similarity 100.0%; Pred. No. 0.65;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CTGAGACCCCTCTCGAGAGTTCCGCT 25
DB 78 CTGAGACCCCTCTCGAGAGTTCCGCT 102

RESULT 4
LOCUS BH638999 534 bp DNA linear GSS 14-FEB-2002
DEFINITION 1008026F08.1EL_x1 1008 - Rescuedu Grid I Zea mays genomic, genomic survey sequence.
ACCESSION BH638999 GI:18663697
VERSION BH638999.1
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
AUTHORS 1 (bases 1 to 534)
JOURNAL Maize genomic sequences found using engineered Rescemu transposon
COMMENT Unpublished (2001)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Very probable ligation site of ends cut by single endonuclease.
Reverse complemented post-ligation sequence from source sequence.
Plate: 108026 row: 10
Class: transposon-tagged.
Location/Qualifiers

FEATURES

source
1..534
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1008 - Rescemu Grid I"
/note="Organ: leaf; Vector: Rescemu (engineered from
pBluescript backbone); Site 1: BamHI; Site 2: BglII;
Rescemu is a 4.9 kb, modified maize mu transposon
designed to allow plasmid rescue from total genomic DNA.
Mu elements insert preferentially into transcription
units. For more information on Rescemu, go to the web
site www.zmdb.iastate.edu and follow the links for
'Rescemu.' Grid I was grown at Berkeley in 2001. DNA was
extracted from leaf punches, double digested using BamHI
and BglII, and ligated to form circular plasmids. DH10B
cells were transformed and then screened on LB plates with
ampicillin."

ORIGIN

Query Match 100.0%; Score 25; DB 8; Length 534;
Best Local Similarity 100.0%; Pred. No. 0.67; 0; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACCCCTCTCGAGATTCGCT 25
|||||
Db 246 CTGACCCCTCTCGAGATTCGCT 270

RESULT 5
LOCUS CG709678 617 bp DNA linear GSS 20-OCT-2003
DEFINITION 1119014D02.2EL.X1 1119 - Rescemu Grid AA Zea mays genomic, genomic
survey sequence.
ACCESSION CG709678
VERSION CG709678.1 GI:37735584
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
AUTHORS 1 (bases 1 to 617)
JOURNAL Maize genomic sequences found using engineered Rescemu transposon
COMMENT Unpublished (2001)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227

FEATURES

source
1..620
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
AUTHORS 1 (bases 1 to 620)
JOURNAL Maize genomic sequences found using engineered Rescemu transposon
COMMENT Unpublished (2001)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Possible ligation site of ends cut by 2 different endonucleases.
Reverse complemented post-ligation sequence from source sequence.
Plate: 1119014 row: 29
Class: transposon-tagged.
Location/Qualifiers

FEATURES

source
1..617
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73/K55"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1119 - Rescemu Grid AA"
/note="Organ: leaf; Vector: Rescemu (engineered from
pBluescript backbone); Site 1: BamHI; Site 2: BglII;
Rescemu is a 4.9 kb, modified maize mu transposon
designed to allow plasmid rescue from total genomic DNA.
Mu elements insert preferentially into transcription
units. For more information on Rescemu, go to the web
site www.zmdb.iastate.edu and follow the links for
'Rescemu.' Grid AA was grown at UC San Diego in 2002. DNA
was extracted from leaf strips, double digested using
BamHI and BglII, and ligated to form circular plasmids.
DH10B cells were transformed and then screened on LB
plates with ampicillin."

ORIGIN

Query Match 100.0%; Score 25; DB 9; Length 617;
Best Local Similarity 100.0%; Pred. No. 0.68; 0; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACCCCTCTCGAGATTCGCT 25
|||||
Db 356 CTGACCCCTCTCGAGATTCGCT 380

RESULT 6
LOCUS CG805054 620 bp DNA linear GSS 10-NOV-2003
DEFINITION 1118056H01.2EL.X1 1118 - Rescemu Grid S Zea mays genomic, genomic
survey sequence.
ACCESSION CG805054
VERSION CG805054.1 GI:38241084
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
AUTHORS 1 (bases 1 to 620)
JOURNAL Maize genomic sequences found using engineered Rescemu transposon
COMMENT Unpublished (2001)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Possible ligation site of ends cut by 2 different endonucleases.
Reverse complemented post-ligation sequence from source sequence.
Plate: 1118056 row: 40
Class: transposon-tagged.
Location/Qualifiers

FEATURES

source
1..620
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"

```

/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1118 - Rescuem Mu Grid S"
/note="Organ: leaf; Vector: Rescuem (engineered from pBluescript backbone); Site 1: BamHI, Site 2: BglII; Rescuem is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on Rescuem, go to the web site 'www.zmblabscat.edu' and follow the links for 'Rescuem'. Grid S was grown at San Diego in 2002. DNA was extracted from leaf strips, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

```

ORIGIN

```

Query Match      100.0%; Score 25; DB 9; Length 620;
Best Local Similarity 100.0%; Pred. No. 0.68;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CTGAGACCCCTCTCGAGAGTTCCGCT 25
Db 358 CTGAGACCCCTCTCGAGAGTTCCGCT 382

```

```

RESULT 7
CO472745      716 bp      mRNA      linear      EST 09-JUN-2004
DEFINITION   GQ0081.BR.1.D19 GQ008: Cambium, phloem and bark of girdled saplings
ACCESSION    CO472745
VERSION       CO472745
KEYWORDS      EST.
SOURCE        Picea glauca (white spruce)
ORGANISM      Picea glauca
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
1 (bases 1 to 716)
Morency,M.-J., Cooke,J., Pavy,N., Parsons,L., Paule,C., Seguin,A.,
Siddiqui,A., Butterfield,Y., Barber,S., Yang,G., Scott,J.,
Arborea EST sequencing in Picea glauca (white spruce)
Unpublished (2004)
Contact: John Mackay
Centre de Recherche en Biologie Forestiere
Universite Laval
Pavillon Charles-Eugene Marchand, Quebec, Quebec, CANADA G1K 7P4
Fax: 418 656 7493
Email: jmackay@rsb.ulaval.ca
Center for Computational Genomics and Bioinformatics (CCGB),
University of Minnesota, MN id identifier: MN519733 Clone ID:
GQ0081.BR.D19 Clones available through: John Mackay, Ph. D.
Professeur adjoint -Assistant professeur EMAIL:
jmackay@rsb.ulaval.ca Centre de Recherche en Biologie Forestiere
(Forest Biology Research Center) Universite Laval Quebec, Quebec
CANADA G1K 7P4
Plate: 1.BR row: 19 column: D
Seq primer: M13 Reverse Primer.

```

FEATURES

source

```

1..716
/organism="Picea glauca"
/mol_type="mRNA"
/strain="Pg-653"
/db_xref="taxon:3330"
/clone="GQ0081.BR.D19"
/sex="Hermaphrodite"
/tissue_type="Vascular cambium, secondary phloem and bark
of trees girdled by removing a ring of bark ca. 1 cm wide
from the midpoint of the main stem"
/dev_stage="Vascular cambium, secondary phloem, and bark
from trees harvested 1 and 7 days after girdling

```

```

treatment"
/lab_host="E. coli DH10B cells"
/clone_lib="GQ008: Cambium, phloem and bark of girdled
saplings"
/note="Organ: Main stem of 4 year old saplings,
approximately 60 cm tall; Vector: pBluescript II SK (+)
XR; Site 1: Eco-RI; Site 2: Xho-I; Tissues were pooled
from above and below the girdle. cDNA was prepared from 5
mg of poly A+ selected RNA and was directionally ligated
into the pBluescript II SK (+) XR vector (Stratagene),
transformed by electroporation into DH10B cells (in
vitrogen) for propagation"

```

ORIGIN

```

Query Match      100.0%; Score 25; DB 7; Length 716;
Best Local Similarity 100.0%; Pred. No. 0.68;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CTGAGACCCCTCTCGAGAGTTCCGCT 25
Db 666 CTGAGACCCCTCTCGAGAGTTCCGCT 690

```

```

RESULT 8
CO472714      765 bp      mRNA      linear      EST 09-JUN-2004
DEFINITION   GQ0081.BR.1.C16 GQ008: Cambium, phloem and bark of girdled saplings
ACCESSION    CO472714
VERSION       CO472714
KEYWORDS      EST.
SOURCE        Picea glauca (white spruce)
ORGANISM      Picea glauca
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
1 (bases 1 to 765)
Morency,M.-J., Cooke,J., Pavy,N., Parsons,L., Paule,C., Seguin,A.,
Siddiqui,A., Butterfield,Y., Barber,S., Yang,G., Scott,J.,
Arborea EST sequencing in Picea glauca (white spruce)
Unpublished (2004)
Contact: John Mackay
Centre de Recherche en Biologie Forestiere
Universite Laval
Pavillon Charles-Eugene Marchand, Quebec, Quebec, CANADA G1K 7P4
Fax: 418 656 7493
Email: jmackay@rsb.ulaval.ca
Center for Computational Genomics and Bioinformatics (CCGB),
University of Minnesota, MN id identifier: MN5197864 Clone ID:
GQ0081.BR.C16 Clones available through: John Mackay, Ph. D.
Professeur adjoint -Assistant professeur EMAIL:
jmackay@rsb.ulaval.ca Centre de Recherche en Biologie Forestiere
(Forest Biology Research Center) Universite Laval Quebec, Quebec
CANADA G1K 7P4
Plate: 1.BR row: 16 column: C
Seq primer: M13 Reverse Primer.

```

FEATURES

source

```

1..765
/organism="Picea glauca"
/mol_type="mRNA"
/strain="Pg-653"
/db_xref="taxon:3330"
/clone="GQ0081.BR.C16"
/sex="Hermaphrodite"
/tissue_type="Vascular cambium, secondary phloem and bark
of trees girdled by removing a ring of bark ca. 1 cm wide
from the midpoint of the main stem"
/dev_stage="Vascular cambium, secondary phloem, and bark
from trees harvested 1 and 7 days after girdling
treatment"
/lab_host="E. coli DH10B cells"
/clone_lib="GQ008: Cambium, phloem and bark of girdled
saplings"

```


/note="Organ: Main stem of 4 year old saplings, approximately 60 cm tall; Vector: pBluescript II SK (+) XR; Site 1: Eco-RI; Site 2: Xho-I; Tissues were pooled from above and below the girdle. cDNA was prepared from 5 mg of poly A+ selected RNA and was directionally ligated into the pBluescript II SK (+) XR vector (Stratagene), transformed by electroporation into DH10B cells (in vitrogen) for propagation"

ORIGIN

Query Match 100.0%; Score 25; DB 7; Length 765;
Best Local Similarity 100.0%; Pred. No. 0.68;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGACCCCTCTCGAGAGTTCCGCT 25
|||||
Db 671 CTGACCCCTCTCGAGAGTTCCGCT 695

RESULT 9
LOCUS CO472736 769 bp mRNA linear EST 09-JUL-2004
DEFINITION GQ0081.BR.1.G06 GQ008: Cambium, phloem and bark of girdled saplings
Picea glauca cDNA clone GQ0081.BR_G06 5', mRNA sequence.
ACCESSION CO472736
VERSION CO472736.1 GI:50141251
KEYWORDS EST.
SOURCE Picea glauca (white spruce)
ORGANISM Picea glauca
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
1 (bases 1 to 769)
Morency,M.-J., Cooke,J., Pavy,N., Parsons,L., Paule,C., Seguin,A., Retzel,E., Butterfield,Y., Barber,S., Yang,G., Stott,J., Siddiqui,A., Holt,R., Maitra,M. and Mackay,J.
Arborea EST sequencing in Picea glauca (white spruce)
Unpublished (2004)
Contact: John Mackay
Centre de Recherche en Biologie Forestiere
Universite Laval
Pavillon Charles-Eugene Marchand, Quebec, CANADA G1K 7P4
Fax: 418 656 7493
Email: jmackay@rsvs.ulaval.ca
Center for Computational Genomics and Bioinformatics (CCGB),
University of Minnesota, MN Id Identifier: MN5198051 Clone ID:
GQ0081.BR.G06 Clones available through: John Mackay, Ph. D.
Professeur adjoint - Assistant professeur EMAIL:
jmackay@rsvs.ulaval.ca Centre de Recherche en Biologie Forestiere
(Forest Biology Research Center) Universite Laval Quebec, Quebec
CANADA G1K 7P4
Plate: 1.BR row: 06 column: G
Seq primer: M13 Reverse Primer.
Location/Qualifiers
1..769

FEATURES

source
/organism="Picea glauca"
/mol_type="mRNA"
/strain="pg-653"
/db_xref="taxon:3330"
/clone="GQ0081.BR.G06"
/sex="Hermaphrodite"
/tissue_type="Vascular cambium, secondary phloem and bark of trees girdled by removing a ring of bark ca. 1 cm wide from the midpoint of the main stem"
/dev_stage="Vascular cambium, secondary phloem, and bark from trees harvested 1 and 7 days after girdling treatment"
/lab_host="E. coli DH10B cells"
/clone_1lb="GQ008: Cambium, phloem and bark of girdled saplings"
/note="Organ: Main stem of 4 year old saplings, approximately 60 cm tall; Vector: pBluescript II SK (+) XR; Site 1: Eco-RI; Site 2: Xho-I; Tissues were pooled from above and below the girdle. cDNA was prepared from 5

mg of poly A+ selected RNA and was directionally ligated into the pBluescript II SK (+) XR vector (Stratagene), transformed by electroporation into DH10B cells (in vitrogen) for propagation"

ORIGIN

Query Match 100.0%; Score 25; DB 7; Length 769;
Best Local Similarity 100.0%; Pred. No. 0.68;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGACCCCTCTCGAGAGTTCCGCT 25
|||||
Db 675 CTGACCCCTCTCGAGAGTTCCGCT 699

RESULT 10
LOCUS CO472678 774 bp mRNA linear EST 09-JUL-2004
DEFINITION GQ0081.BR.1.D17 GQ008: Cambium, phloem and bark of girdled saplings
Picea glauca cDNA clone GQ0081.BR_D17 5', mRNA sequence.
ACCESSION CO472678
VERSION CO472678.1 GI:50141128
KEYWORDS EST.
SOURCE Picea glauca (white spruce)
ORGANISM Picea glauca
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
1 (bases 1 to 774)
Morency,M.-J., Cooke,J., Pavy,N., Parsons,L., Paule,C., Seguin,A., Retzel,E., Butterfield,Y., Barber,S., Yang,G., Stott,J., Siddiqui,A., Holt,R., Maitra,M. and Mackay,J.
Arborea EST sequencing in Picea glauca (white spruce)
Unpublished (2004)
Contact: John Mackay
Centre de Recherche en Biologie Forestiere
Universite Laval
Pavillon Charles-Eugene Marchand, Quebec, CANADA G1K 7P4
Fax: 418 656 7493
Email: jmackay@rsvs.ulaval.ca
Center for Computational Genomics and Bioinformatics (CCGB),
University of Minnesota, MN Id Identifier: MN5197337 Clone ID:
GQ0081.BR.D17 Clones available through: John Mackay, Ph. D.
Professeur adjoint - Assistant professeur EMAIL:
jmackay@rsvs.ulaval.ca Centre de Recherche en Biologie Forestiere
(Forest Biology Research Center) Universite Laval Quebec, Quebec
CANADA G1K 7P4
Plate: 1.BR row: 17 column: D
Seq primer: M13 Reverse Primer.
Location/Qualifiers
1..774

FEATURES

source
/organism="Picea glauca"
/mol_type="mRNA"
/strain="pg-653"
/db_xref="taxon:3330"
/clone="GQ0081.BR.D17"
/sex="Hermaphrodite"
/tissue_type="Vascular cambium, secondary phloem and bark of trees girdled by removing a ring of bark ca. 1 cm wide from the midpoint of the main stem"
/dev_stage="Vascular cambium, secondary phloem, and bark from trees harvested 1 and 7 days after girdling treatment"
/lab_host="E. coli DH10B cells"
/clone_1lb="GQ008: Cambium, phloem and bark of girdled saplings"
/note="Organ: Main stem of 4 year old saplings, approximately 60 cm tall; Vector: pBluescript II SK (+) XR; Site 1: Eco-RI; Site 2: Xho-I; Tissues were pooled from above and below the girdle. cDNA was prepared from 5 mg of poly A+ selected RNA and was directionally ligated into the pBluescript II SK (+) XR vector (Stratagene), transformed by electroporation into DH10B cells (in vitrogen) for propagation"

DB 687 CTGACCCCTCTCGAGAGTTCGCT 711

RESULT 13

LOCUS CO472694

DEFINITION GQ0081.BR.1 J17 GQ008: Cambium, phloem and bark of girdled saplings

ACCESSION Picea glauca cDNA clone GQ0081.BR_J17 5', mRNA sequence.

CO472694

VERSION CO472694.1 GI:50141166

KEYWORDS EST.

SOURCE Picea glauca (white spruce)

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.

REFERENCE 1 (bases 1 to 819)

AUTHORS Morency,M.-J., Cooke,J., Pavy,N., Parsons,L., Paule,C., Seguin,A., Retzel,E., Butterfield,Y., Barber,S., Yang,G., Stott,J., Siddiqui,A., Holt,R., Marra,M. and Mackay,J.

TITLE Arborea EST sequencing in Picea glauca (white spruce)

JOURNAL Unpublished (2004)

COMMENT Contact: John Mackay
Centre de Recherche en Biologie Forestiere
Universite Laval
Pavillon Charles-Eugene Marchand, Quebec, Quebec, CANADA G1K 7P4
Fax: 418 656 7493
Email: jmackay@rsvs.ulaval.ca
Center for Computational Genomics and Bioinformatics (CCGB),
University of Minnesota, MN id Identifier: MN5197649 Clone ID:
GQ0081.BR.J17 Clones available through: John Mackay, Ph. D.
Professeur adjoint -Assistant professor EMAIL:
jmackay@rsvs.ulaval.ca Centre de Recherche en Biologie Forestiere
(Forest Biology Research Center) Universite Laval Quebec, Quebec
CANADA G1K 7P4
Plate: 1.BR row: 17 column: J
Seq primer: M3 Reverse Primer.

FEATURES

source

1. 819

Location/Qualifiers

/organism="Picea glauca"

/mol_type="mRNA"

/strain="pg-653"

/db_xref="taxon:3330"

/clone="GQ0081.BR.J17"

/sex="Hermaphrodite"

/tissue_type="Vascular cambium, secondary phloem and bark of trees girdled by removing a ring of bark ca. 1 cm wide from the midpoint of the main stem"

/dev_stage="Vascular cambium, secondary phloem, and bark from trees harvested 1 and 7 days after girdling treatment"

/lab_host="E. coli DH10B cells"

/clone_1lb="GQ008: Cambium, phloem and bark of girdled saplings"

/note="Organ: Main stem of 4 year old saplings, approximately 60 cm tall; Vector: pBluescript II SK (+) XR; Site_1: Eco-RI; Site_2: Xho-I; Tissues were pooled from above and below the girdle. cDNA was prepared from 5 mg of poly A+ selected RNA and was directionally ligated into the pBluescript II SK (+) XR vector (Stratagene), transformed by electroporation into DH10B cells (in vitrogen) for propagation"

ORIGIN

Query Match 100.0%; Score 25; DB 7; Length 819;
Best Local Similarity 100.0%; Pred. No. 0.69;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGACCCCTCTCGAGAGTTCGCT 25

DB 695 CTGACCCCTCTCGAGAGTTCGCT 719

RESULT 14

CO472674

LOCUS CO472674

DEFINITION GQ0081.BR.1 B11 GQ008: Cambium, phloem and bark of girdled saplings

ACCESSION Picea glauca cDNA clone GQ0081.BR_B11 5', mRNA sequence.

CO472674

VERSION CO472674.1 GI:50141124

KEYWORDS EST.

SOURCE Picea glauca (white spruce)

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.

REFERENCE 1 (bases 1 to 820)

AUTHORS Morency,M.-J., Cooke,J., Pavy,N., Parsons,L., Paule,C., Seguin,A., Retzel,E., Butterfield,Y., Barber,S., Yang,G., Stott,J., Siddiqui,A., Holt,R., Marra,M. and Mackay,J.

TITLE Arborea EST sequencing in Picea glauca (white spruce)

JOURNAL Unpublished (2004)

COMMENT Contact: John Mackay
Centre de Recherche en Biologie Forestiere
Universite Laval
Pavillon Charles-Eugene Marchand, Quebec, Quebec, CANADA G1K 7P4
Fax: 418 656 7493
Email: jmackay@rsvs.ulaval.ca
Center for Computational Genomics and Bioinformatics (CCGB),
University of Minnesota, MN id Identifier: MN5197207 Clone ID:
GQ0081.BR.B11 Clones available through: John Mackay, Ph. D.
Professeur adjoint -Assistant professor EMAIL:
jmackay@rsvs.ulaval.ca Centre de Recherche en Biologie Forestiere
(Forest Biology Research Center) Universite Laval Quebec, Quebec
CANADA G1K 7P4
Plate: 1.BR row: 11 column: B
Seq primer: M3 Reverse Primer.

FEATURES

source

1. 820

Location/Qualifiers

/organism="Picea glauca"

/mol_type="mRNA"

/strain="pg-653"

/db_xref="taxon:3330"

/clone="GQ0081.BR.B11"

/sex="Hermaphrodite"

/tissue_type="Vascular cambium, secondary phloem and bark of trees girdled by removing a ring of bark ca. 1 cm wide from the midpoint of the main stem"

/dev_stage="Vascular cambium, secondary phloem, and bark from trees harvested 1 and 7 days after girdling treatment"

/lab_host="E. coli DH10B cells"

/clone_1lb="GQ008: Cambium, phloem and bark of girdled saplings"

/note="Organ: Main stem of 4 year old saplings, approximately 60 cm tall; Vector: pBluescript II SK (+) XR; Site_1: Eco-RI; Site_2: Xho-I; Tissues were pooled from above and below the girdle. cDNA was prepared from 5 mg of poly A+ selected RNA and was directionally ligated into the pBluescript II SK (+) XR vector (Stratagene), transformed by electroporation into DH10B cells (in vitrogen) for propagation"

ORIGIN

Query Match 100.0%; Score 25; DB 7; Length 820;
Best Local Similarity 100.0%; Pred. No. 0.69;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGACCCCTCTCGAGAGTTCGCT 25

DB 695 CTGACCCCTCTCGAGAGTTCGCT 719

RESULT 15

LOCUS CO472692

DEFINITION GQ0081.BR.1 J08 GQ008: Cambium, phloem and bark of girdled saplings

ACCESSION Picea glauca cDNA clone GQ0081.BR_J08 5', mRNA sequence.

CO472692

VERSION CO472692.1 GI:50141124

KEYWORDS EST.

SOURCE Picea glauca (white spruce)

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.

REFERENCE 1 (bases 1 to 825)

AUTHORS Morency,M.-J., Cooke,J., Pavy,N., Parsons,L., Paule,C., Seguin,A., Retzel,E., Butterfield,Y., Barber,S., Yang,G., Stott,J., Siddiqui,A., Holt,R., Marra,M. and Mackay,J.

TITLE Arborea EST sequencing in Picea glauca (white spruce)

JOURNAL Unpublished (2004)

COMMENT Contact: John Mackay
Centre de Recherche en Biologie Forestiere
Universite Laval
Pavillon Charles-Eugene Marchand, Quebec, Quebec, CANADA G1K 7P4
Fax: 418 656 7493
Email: jmackay@rsvs.ulaval.ca
Center for Computational Genomics and Bioinformatics (CCGB),
University of Minnesota, MN id Identifier: MN5197207 Clone ID:
GQ0081.BR.J08 Clones available through: John Mackay, Ph. D.
Professeur adjoint -Assistant professor EMAIL:
jmackay@rsvs.ulaval.ca Centre de Recherche en Biologie Forestiere
(Forest Biology Research Center) Universite Laval Quebec, Quebec
CANADA G1K 7P4
Plate: 1.BR row: 11 column: B
Seq primer: M3 Reverse Primer.

FEATURES

source

1. 825

Location/Qualifiers

/organism="Picea glauca"

/mol_type="mRNA"

/strain="pg-653"

/db_xref="taxon:3330"

/clone="GQ0081.BR.B11"

/sex="Hermaphrodite"

/tissue_type="Vascular cambium, secondary phloem and bark of trees girdled by removing a ring of bark ca. 1 cm wide from the midpoint of the main stem"

/dev_stage="Vascular cambium, secondary phloem, and bark from trees harvested 1 and 7 days after girdling treatment"

/lab_host="E. coli DH10B cells"

/clone_1lb="GQ008: Cambium, phloem and bark of girdled saplings"

/note="Organ: Main stem of 4 year old saplings, approximately 60 cm tall; Vector: pBluescript II SK (+) XR; Site_1: Eco-RI; Site_2: Xho-I; Tissues were pooled from above and below the girdle. cDNA was prepared from 5 mg of poly A+ selected RNA and was directionally ligated into the pBluescript II SK (+) XR vector (Stratagene), transformed by electroporation into DH10B cells (in vitrogen) for propagation"

ORIGIN

Query Match 100.0%; Score 25; DB 7; Length 820;
Best Local Similarity 100.0%; Pred. No. 0.69;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGACCCCTCTCGAGAGTTCGCT 25

DB 695 CTGACCCCTCTCGAGAGTTCGCT 719

RESULT 15

ACCESSION CO472692
 VERSION CO472692.1 GI:50141162
 KEYWORDS
 SOURCE
 ORGANISM
 Picea glauca (white spruce)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
 1 (bases 1 to 825)
 Moroney,M.-J., Cooke,J., Pavy,N., Parsons,L., Paule,C., Seguin,A.,
 Retzel,E., Butterfield,Y., Barber,S., Yang,G., Scott,J.,
 Siddiqui,A., Holt,R., Marra,M. and Mackay,J.
 Arborea EST sequencing in Picea glauca (white spruce)
 Unpublished (2004)
 Contact: John Mackay
 Centre de Recherche en Biologie Forestiere
 Universite Laval
 Pavillon Charles-Eugene Marchand, Quebec, Quebec, CANADA G1K 7P4
 Fax: 418 656 7493
 Email: jmackay@svs.ulaval.ca
 Center for Computational Genomics and Bioinformatics (CCGB),
 University of Minnesota, MN id Identifier: MN5197621 Clone ID:
 GQ0081.BR_J08 Clones available through: John Mackay, Ph. D.
 Professeur adjoint -Assistant professor EMAIL:
 jmackay@svs.ulaval.ca Centre de Recherche en Biologie Forestiere
 (Forest Biology Research Center) Universite Laval Quebec, Quebec
 CANADA G1K 7P4
 Plate: 1.BR row: 08 column: J
 Seq primer: M13 Reverse primer.
 Location/Qualifiers

FEATURES

Source

1..825
 /organism="Picea glauca"
 /mol_type="mRNA"
 /strain="PG-553"
 /db_xref="taxon:3330"
 /clone="GQ0081.BR_J08"
 /sex="Hermaphrodite"
 /tissue_type="Vascular cambium, secondary phloem and bark
 of trees girdled by removing a ring of bark ca. 1 cm wide
 from the midpoint of the main stem"
 /dev_stage="Vascular cambium, secondary phloem, and bark
 from trees harvested 1 and 7 days after girdling
 treatment"
 /lab_host="E. coli DH10B cells"
 /clone_id="GQ0081 Cambium, phloem and bark of girdled
 saplings"
 /note="Organ: Main stem of 4 year old saplings,
 approximately 60 cm tall; Vector: pBluescript II SK (+)
 XR; Site_1: Eco-RI; Site_2: Xho-I; Tissues were pooled
 from above and below the girdle. cDNA was prepared from 5
 mg of poly A+ selected RNA and was directionally ligated
 into the pBluescript II SK (+) XR vector (Stratagene),
 transformed by electroporation into DH10B cells (in
 vitro) for propagation"

ORIGIN

Query Match 100.0%; Score 25; DB 7; Length 825;
 Best Local Similarity 100.0%; Pred. No. 0.69;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CTGAGCCCTCTCTCGAGAGTTCCGCT 25
 |||||
 DB 695 CTGAGCCCTCTCTCGAGAGTTCCGCT 719
 |||||

Search completed: March 1, 2005, 01:58:18
 Job time : 1738.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2005, 23:26:41 ; Search time 192 Seconds
(without alignments)
678.304 Million cell updates/sec

Title: US-10-086-062-5

Perfect score: 22

Sequence: 1 gacacgtagatgagtcacac 22

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : N_Geneseq16Dec04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001s:*
- 5: geneseqn2002s:*
- 6: geneseqn2003s:*
- 7: geneseqn2004s:*
- 8: geneseqn2005s:*
- 9: geneseqn2006s:*
- 10: geneseqn2007s:*
- 11: geneseqn2008s:*
- 12: geneseqn2009s:*
- 13: geneseqn2010s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	22	AAH42764	AAH42764 A promote
2	22	100.0	22	ADD24070	ADD24070 Pea lecti
3	18.8	85.5	110000	ABD32968_1	Continuation (2 of
4	17.2	78.2	142	AAI20672	AAI20672 Probe #10
5	17.2	78.2	142	ABA65716	ABA65716 Human foe
6	17.2	78.2	142	AAI45884	AAI45884 Probe #14
7	17.2	78.2	142	ABA32803	ABA32803 Probe #11
8	17.2	78.2	142	AAK39662	AAK39662 Human bon
9	17.2	78.2	142	AAK14121	AAK14121 Human bra
10	17.2	78.2	142	ABK39453	ABK39453 Human liv
11	17.2	78.2	142	ABK13960	ABK13960 Human gen
12	17.2	78.2	142	ABA53123	ABA53123 Human foe
13	17.2	78.2	142	AAI32730	AAI32730 Probe #13
14	17.2	78.2	142	ABA22897	ABA22897 Probe #13
15	17.2	78.2	142	AAK26828	AAK26828 Human don
16	17.2	78.2	142	AAK01373	AAK01373 Human bra
17	17.2	78.2	142	ABK26422	ABK26422 Human liv
18	17.2	78.2	142	ABK01426	ABK01426 Human gen
19	17.2	78.2	142	AAK26633	AAK26633 Human gen
20	17.2	78.2	142	AAK26633	AAK26633 Human gen

ALIGNMENTS

21	17.2	78.2	714	4	AAK26632	AAK26632 Human gen
22	17.2	78.2	714	4	AAK86129	AAK86129 Human imm
23	17.2	78.2	714	4	AAK86128	AAK86128 Human imm
24	17.2	78.2	714	4	AAI36601	AAI36601 Human mus
25	17.2	78.2	714	4	AAI36602	AAI36602 Human mus
26	17.2	78.2	714	8	ABK73982	ABK73982 Human nov
27	17.2	78.2	714	8	ABK73981	ABK73981 Human nov
28	17.2	78.2	714	8	ABK59589	ABK59589 CDNA enco
29	17.2	78.2	714	8	ABK59590	ABK59590 CDNA enco
30	17.2	78.2	714	12	ADJ30339	ADJ30339 Human mus
31	17.2	78.2	714	12	ADJ30340	ADJ30340 Human mus
32	17.2	78.2	110000	11	ADM27081_07	Continuation (8 of
33	17.2	78.2	163321	11	ACN43898	ACN43898 Human gen
34	16.8	76.4	6469	2	AAQ11690	AAQ11690 Entire se
35	16.2	73.6	555	10	ADK59966	ADK59966 Plant DNA
36	16.2	73.6	583	13	ADQ48845	ADQ48845 Novel can
37	16.2	73.6	1644	13	ADR93347	ADR93347 Novel S.
38	16.2	73.6	1980	12	ADM91890	ADM91890 S. pneumo
39	16.2	73.6	2250	10	ABX06397	ABX06397 S. pneumo
40	16.2	73.6	2253	4	AAK55670	AAK55670 Streptoco
41	16.2	73.6	2253	8	ACN49821	ACN49821 Prokaryot
42	16.2	73.6	2553	8	ADN70404	ADN70404 Rice gene
43	16.2	73.6	3234	4	ABJ30454	ABJ30454 Drosophi1
44	16.2	73.6	3340	4	ABJ19058	ABJ19058 Drosophi1
45	16.2	73.6	3421	4	ABJ26860	ABJ26860 Drosophi1

RESULT 1

AAH42764 standard; DNA; 22 BP.

AAH42764;

01-OCT-2001 (first entry)

A promoter element or transcription binding site.

Promoter element; transcription binding site; plant promoter; SMER;

KW synthetic multimeric promoter element region; gene expression;

KX insect resistance; herbicide resistance; ss.

OS Pisum sativum.

PN WO200153476-A2.

PD 26-JUL-2001.

PF 19-JAN-2001; 2001WO-US002024.

PR 21-JAN-2000; 2000US-0177437P.

PA (PION-) PIONEER HI-BRED INT INC.

PI Bruce WB, Niu X;

XX WPI; 2001-476118/51.

XX New plant promoters with synthetic multimeric promoter element regions,

XX useful in plant molecular biology, particularly in regulating gene

XX expression in plants to increase resistance against insects or

XX herbicides.

XX Example 1; Fig 1; 67bp; English.

XX AAH42709-72 represent promoter elements or transcription binding sites.

XX They are used to construct synthetic multimeric promoter element

XX regions (SMERs). The specification describes plant promoters which

XX comprise SMERs. The plant promoters are useful in plant molecular

XX biology, particularly in regulating gene expression in plants. The

XX promoters are especially useful for transforming plants or plant cells,

CC e.g. to increase resistance against insects or herbicides
XX
SQ Sequence 22 BP; 8 A; 5 C; 5 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 4; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACGTAGATGATCATCAC 22
DB 1 GACACGTAGATGATCATCAC 22

RESULT 2
AADD24070
ID AAD24070 standard; DNA; 22 BP.

XX AAD24070;
XX
DT 09-APR-2002 (first entry)
DE Pea lectin gene Psi element.

XX Gene expression; maize; ubiquitin promoter; Ubi-1; HSE;
KM heat shock element; agronomic gene; pea; lectin gene; Psi element; ds.
XX

OS Pisum sativum.

PN WO200194394-A2.

PD 13-DEC-2001.

PF 08-JUN-2001; 2001WO-US018689.

PR 09-JUN-2000; 2000US-00590558.

XX (PROD-) PRODIGENE INC.

PI Jilka JM, Hood BE, Howard JA;

DR WPI; 2002-122117/16.

PT New promoter sequences for causing expression of a structural gene
PT especially agronomic gene or open reading frame in a plant cell,
XX comprises engineered versions of the maize ubiquitin promoter.

PS Claim 11; Page 55; 68pp; English.

CC The invention relates to a promoter sequence capable of directing
CC expression of a nucleotide sequence in a plant cell comprising maize
CC ubiquitin (Ubi-1) promoter sequence with a modification so that it does
CC not include two overlapping heat shock elements (HSE) or it directs
CC expression to increase the endosperm/embryo expression ratio of the
CC protein when compared to the ratio from a wild-type ubiquitin promoter.
CC The modified Ubi-1 promoter comprises a deletion of 3', 5' or both HSEs,
CC two non-overlapping/adjacent HSEs, replacement of HSEs with a trimer of a
CC seed specific element from the promoter of pea lectin gene Psi, or
CC insertion of a transcription factor binding site in the HSE region. An
CC expression construct comprising modified Ubi-1 promoter is useful for
CC causing expression of a structural gene (agronomic genes) or open reading
CC frame in a plant cell. The modified Ubi-1 promoter increases expression
CC levels beyond those observed with native ubiquitin promoter. The present
CC sequence is pea lectin gene Psi element used in the promoter of the
CC invention
XX

SQ Sequence 22 BP; 8 A; 5 C; 5 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.24;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACGTAGATGATCATCAC 22
|||||

DB 1 GACACGTAGATGATCATCAC 22

RESULT 3
ABD32968_1
Continuation (2 of 8) of ABD32968 from base 100001 (Human cancer-associated genomic DNA)
WP Sequence split into 8 fragments LOCUS ABD32968 Accession Abd32968

WP	Fragment Name	Begin	End
WP	ABD32968_0	1	110000
WP	ABD32968_1	100001	210000
WP	ABD32968_2	200001	310000
WP	ABD32968_3	300001	410000
WP	ABD32968_4	400001	510000
WP	ABD32968_5	500001	610000
WP	ABD32968_6	600001	710000
WP	ABD32968_7	700001	779603

Query Match 85.5%; Score 18.8; DB 13; Length 110000;
Best Local Similarity 90.9%; Pred. No. 29;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GACACGTAGATGATCATCAC 22
DB 42953 GACACGTAGATGATCATCAC 42954

RESULT 4
AAI20672
ID AAI20672 standard; DNA; 142 BP.

AC AAI20672;

DT 12-OCT-2001 (first entry)

DE Probe #10605 for gene expression analysis in human cervical cell sample.

XX Probe; human; microarray; gene expression; cervical epithelial cell;

XX cervical cancer; ss.

XX Homo sapiens.

OS WO200157278-A2.

PN 30-JAN-2001; 2001WO-US000670.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488901/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human cervical epithelial cells.

PS Claim 25; SEQ ID NO 10605; 487bp; English.

CC The present invention relates to human single exon nucleic acid probes
CC (SNP). The present sequence is one such probe. The SNPs are derived
CC from human Hela cells. The SNPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging of
CC diseases of the cervix, notably cervical cancer. Note: The sequence data

CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 142 BP; 52 A; 40 C; 20 G; 30 T; 0 U; 0 Other;

Query Match 78.2%; Score 17.2; DB 4; Length 142;

Best Local Similarity 86.4%; Pred. No. 86;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 GACACGTAGATGAGTCATCAC 22

DB 80 GACACATAAATCAGTCATCAC 101

RESULT 5

ABA65716
ID ABA65716 standard; DNA; 142 BP.

AC ABA65716;

DT 01-FEB-2002 (first entry)

XX Human foetal liver single exon nucleic acid probe #14021.

XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

OS Homo sapiens.

PN WO200157277-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000669.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483447/52.

PT Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human fetal liver.

PS Claim 4; SEQ ID NO 14021; 639pp + Sequence Listing; English.

CC The invention relates to a single exon nucleic acid probe for measuring

CC human gene expression in a sample derived from human foetal liver. The

CC single exon nucleic acid probes may be used for predicting, measuring and

CC displaying gene expression in samples derived from human fetal liver. The

CC present sequence is a single exon nucleic acid probe of the invention.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 142 BP; 52 A; 40 C; 20 G; 30 T; 0 U; 0 Other;

Query Match 78.2%; Score 17.2; DB 4; Length 142;

Best Local Similarity 86.4%; Pred. No. 86;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 GACACGTAGATGAGTCATCAC 22

DB 80 GACACATAAATCAGTCATCAC 101

RESULT 6
AA145884
ID AA145884 standard; DNA; 142 BP.

XX AA145884;

DT 17-OCT-2001 (first entry)

XX Probe #14570 used to measure gene expression in human placenta sample.

XX Probe; microarray; human; placenta; antenatal diagnosis;

XX genetic disorder; ss.

OS Homo sapiens.

PN WO200157272-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000663.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488897/53.

PT Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human placenta.

PS Claim 25; SEQ ID NO 14570; 654pp; English.

CC The present invention relates to single exon nucleic acid probes (SENP).

CC The present sequence is one such probe. The probes are useful for

CC producing a microarray for predicting, measuring and displaying gene

CC expression in samples derived from human placenta. The probes are useful

CC for antenatal diagnosis of human genetic disorders

XX Sequence 142 BP; 52 A; 40 C; 20 G; 30 T; 0 U; 0 Other;

Query Match 78.2%; Score 17.2; DB 4; Length 142;

Best Local Similarity 86.4%; Pred. No. 86;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 GACACGTAGATGAGTCATCAC 22

DB 80 GACACATAAATCAGTCATCAC 101

RESULT 7

ABA32803
ID ABA32803 standard; DNA; 142 BP.

AC ABA32803;

DT 23-JAN-2002 (first entry)

XX Probe #11269 for gene expression analysis in human heart cell sample.

XX Human; gene expression; heart; microarray; vascular system; probe;

XX cardiovascular disease; hypertension; cardiac arrhythmia;

XX congenital heart disease; ss.

OS Homo sapiens.

```

XX PN WO200157274-A2.
XX XX
XX PD 09-AUG-2001.
XX XX
XX PF 30-JAN-2001; 2001WO-US000666.
XX XX
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-0207456P.
XX PR 03-AUG-2000; 2000US-00608408.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488990/53.
XX DR
XX PT Single exon nucleic acid probes for analyzing gene expression in human
XX PT hearts.
XX PS Claim 4; SEQ ID NO 11269; 530bp; English.
XX XX
XX CC The present invention relates to single exon nucleic acid probes for
XX CC measuring human gene expression in a sample derived from human heart. The
XX CC present sequence is one such probe. The probes may be used for
XX CC predicting, measuring and displaying gene expression in samples derived
XX CC from the human heart via microarrays. By measuring gene expression, the
XX CC probes are useful for predicting, diagnosing, grading, staging,
XX CC monitoring and prognosing diseases of the human heart and vascular system
XX CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
XX CC congenital heart disease. Note: The sequence data for this patent did not
XX CC form part of the printed specification, but was obtained in electronic
XX CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ
XX Sequence 142 BP; 52 A; 40 C; 20 G; 30 T; 0 U; 0 Other;
XX
XX Query Match 78.2%; Score 17.2; DB 4; Length 142;
XX Best Local Similarity 86.4%; Pred. No. 86;
XX Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX Oy 1 GACACGTAGATGATCATCACC 22
XX Db 80 GACACATTAATCATCATCACC 101
XX
XX RESULT 8
XX AAK39862
XX ID AAK39862 standard; DNA; 142 BP.
XX XX
XX AC AAK39862;
XX XX
XX DT 06-NOV-2001 (first entry)
XX XX
XX DE Human bone marrow expressed single exon probe SEQ ID NO: 14419.
XX XX
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;
XX KW microarray; cancer; leukemia; lymphoma; myeloma; ss.
XX OS Homo sapiens.
XX XX
XX PN WO200157276-A2.
XX PD
XX PD 09-AUG-2001.
XX PF
XX PF 30-JAN-2001; 2001WO-US000668.
XX PR
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.

```

```

PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX XX
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488990/53.
XX DR
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human bone marrow.
XX PS Example 4; SEQ ID NO 14419; 658bp + Sequence Listing; English.
XX XX
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC bone marrow. They can be used to measure gene expression in bone marrow
XX CC samples, which may enable the improved diagnosis and treatment of cancers
XX CC such as lymphoma, leukemia and myeloma. The present sequence is one of
XX CC the probes of the invention
XX SQ
XX Sequence 142 BP; 52 A; 40 C; 20 G; 30 T; 0 U; 0 Other;
XX
XX Query Match 78.2%; Score 17.2; DB 4; Length 142;
XX Best Local Similarity 86.4%; Pred. No. 86;
XX Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX Oy 1 GACACGTAGATGATCATCACC 22
XX Db 80 GACACATTAATCATCATCACC 101
XX
XX RESULT 9
XX AAK14121
XX ID AAK14121 standard; DNA; 142 BP.
XX XX
XX AC AAK14121;
XX XX
XX DT 05-NOV-2001 (first entry)
XX XX
XX DE Human brain expressed single exon probe SEQ ID NO: 14112.
XX XX
XX KW Human; brain expressed exon; gene expression analysis; probe; microarray;
XX KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
XX KW ss.
XX OS Homo sapiens.
XX XX
XX PN WO200157275-A2.
XX PD
XX PD 09-AUG-2001.
XX PF
XX PF 30-JAN-2001; 2001WO-US000667.
XX PR
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-483446/52.
XX DR
XX PT Single exon nucleic acid probes for analyzing gene expression in human
XX PT brains.
XX

```


PS Example 4; SEQ ID NO 14112; 650bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention
XX
SQ Sequence 142 BP; 52 A; 40 C; 20 G; 30 T; 0 U; 0 Other;
XX
Query Match 78.2%; Score 17.2; DB 4; Length 142;
Best Local Similarity 86.4%; Pred. No. 86;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
OY 1 GACACGTAGATGAGTCATCAC 22
DB 80 GACACATAAATCAGTCATCAC 101
XX
RESULT 10
ABS39453
ID ABS39453 standard; DNA; 142 BP.
XX
AC ABS39453;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver single exon probe, SEQ ID NO 14443.
XX
KW Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000664.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488898/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human adult liver.
XX
PS Claim 4; SEQ ID NO 14443; 658bp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABS25011-ABS51005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The

CC sequence information for this patent does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 142 BP; 52 A; 40 C; 20 G; 30 T; 0 U; 0 Other;
XX
Query Match 78.2%; Score 17.2; DB 4; Length 142;
Best Local Similarity 86.4%; Pred. No. 86;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
OY 1 GACACGTAGATGAGTCATCAC 22
DB 80 GACACATAAATCAGTCATCAC 101
XX
RESULT 11
ABS13960
ID ABS13960 standard; DNA; 142 BP.
XX
AC ABS13960;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe ORF from lung SEQ ID NO 13951.
XX
KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples.
XX
PS Claim 4; SEQ ID NO 13951; 634bp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 1387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the

CC array; identifying exons in a eukaryotic genome, comprising (a)
 CC algorithmically predicting at least one exon from genomic sequences of
 CC the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene expression
 CC analysis, and for identifying exons in a gene, particularly using human
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
 CC Karsenger syndrome, fibrocystic pulmonary dysplasia, primary ciliary
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
 CC present sequence is a single exon probe open reading frame of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 142 BP; 52 A; 40 C; 20 G; 30 T; 0 U; 0 Other;
 Query Match 78.2%; Score 17.2; DB 6; Length 142;
 Best Local Similarity 86.4%; Pred. No. 86;
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GACACGTAGAGTGCATCAC 22
 |||||
 DB 80 GACACATTAATCATCATCAC 101

RESULT 12

AA11458
 ID AA11458 standard; DNA; 483 BP.

AC AA11458;

DT 12-OCT-2001 (first entry)

DE Probe #1391 for gene expression analysis in human cervical cell sample.

KW Probe; human; microarray; gene expression; cervical epithelial cell;
 KM cervical cancer; ss.

OS Homo sapiens.

PN WO200157278-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000670.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI, 2001-488901/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human cervical epithelial cells.
 PS Claim 25; SEQ ID NO 1391; 487bp; English.

CC The present invention relates to human single exon nucleic acid probes
 CC (SENPs). The present sequence is one such probe. The SENPs are derived
 CC from human Hela cells. The SENPs can be used to produce a single exon
 CC microarray, which can be used for measuring human gene expression in a
 CC sample derived from human cervical epithelial cells. By measuring gene
 CC expression, the probes are therefore useful in grading and/or staging of
 CC diseases of the cervix, notably cervical cancer. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 483 BP; 151 A; 106 C; 99 G; 127 T; 0 U; 0 Other;

Query Match 78.2%; Score 17.2; DB 4; Length 483;
 Best Local Similarity 86.4%; Pred. No. 1e+02;
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GACACGTAGAGTGCATCAC 22
 |||||
 DB 402 GACACATTAATCATCATCAC 423

RESULT 13

ABA53123
 ID ABA53123 standard; DNA; 483 BP.

AC ABA53123;

DT 01-FEB-2002 (first entry)

DE Human foetal liver single exon nucleic acid probe #1428.

KM Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

OS Homo sapiens.

PN WO200157277-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000669.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI, 2001-483447/52.

PT Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human foetal liver.

PS Claim 1; SEQ ID NO 1428; 639bp + Sequence Listing; English.

CC The invention relates to a single exon nucleic acid probe for measuring
 CC human gene expression in a sample derived from human foetal liver. The
 CC single exon nucleic acid probes may be used for predicting, measuring and
 CC displaying gene expression in samples derived from human foetal liver. The
 CC present sequence is a single exon nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 483 BP; 151 A; 106 C; 99 G; 127 T; 0 U; 0 Other;

Query Match 78.2%; Score 17.2; DB 4; Length 483;
Best Local Similarity 86.4%; Pred. No. 1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 GACACGTAGATGACTCATCAC 22
DB 402 GACACATATAATCAGTCATCAC 423

RESULT 14

AA132730
ID AA132730 standard; DNA; 483 BP.

AA132730;

DT 17-OCT-2001 (first entry)

DE Probe #1416 used to measure gene expression in human placenta sample.

XX Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder; ss.

OS Homo sapiens.

PN WO200157272-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000663.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

(MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488897/53.

PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human placenta.

PS Claim 25; SEQ ID NO 1416; 654bp; English.

CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders

XX
XX Sequence 483 BP; 151 A; 106 C; 99 G; 127 T; 0 U; 0 Other;

Query Match 78.2%; Score 17.2; DB 4; Length 483;
Best Local Similarity 86.4%; Pred. No. 1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 GACACGTAGATGACTCATCAC 22
DB 402 GACACATATAATCAGTCATCAC 423

RESULT 15
ABA22897

ID ABA22897 standard; DNA; 483 BP.

XX ABA22897;

DT 23-JAN-2002 (first entry)

DE Probe #1363 for gene expression analysis in human heart cell sample.

XX Human; gene expression; heart; microarray; vascular system; probe;

XX cardiovascular disease; hypertension; cardiac arrhythmia;

XX congenital heart disease; ss.

OS Homo sapiens.

PN WO200157274-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000666.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

(MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488897/53.

PT Single exon nucleic acid probes for analyzing gene expression in human
XX hearts.

PS Claim 1; SEQ ID NO 1363; 530bp; English.

CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX
XX Sequence 483 BP; 151 A; 106 C; 99 G; 127 T; 0 U; 0 Other;

Query Match 78.2%; Score 17.2; DB 4; Length 483;
Best Local Similarity 86.4%; Pred. No. 1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 GACACGTAGATGACTCATCAC 22
DB 402 GACACATATAATCAGTCATCAC 423

Search completed: February 28, 2005, 23:51:50
Job time : 198 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2005, 23:32:36 ; Search time 641.714 Seconds
(without alignments)
1661.198 Million cell updates/sec

Title: US-10-086-062-5
Perfect score: 22
Sequence: 1 gacacgtagatgagtcacac 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	22	6 AX207103	AX207103 Sequence
2	22	100.0	22	6 AX358110	AX358110 Sequence
3	22	100.0	827	12 A1192160	A1192160 Synthetic
4	22	100.0	1659	8 PEALCEB	M18160 Pea PSL1 ge
5	22	100.0	1701	6 A14134	A14134 promoter an
6	22	100.0	1701	8 PSLECA	Y00440 Pismu sativ
7	22	100.0	3360	8 PSLECTIN	X66368 P. sativum p
8	18.8	85.5	136492	9 AC011143	AC011143 Homo sapi
9	18.8	85.5	195156	9 AC090660	AC090660 Homo sapi
10	18.8	85.5	198127	2 AP001593	AP001593 Homo sapi
11	18.8	85.5	349980	6 CQ870293	CQ870293 Sequence
12	18.4	83.6	226810	2 AC125804	AC125804 Rattus no
13	18	81.8	146407	5 BX927258	BX927258 zebrafish
14	17.8	80.9	5322	5 CHROVALSA	M29020 Chicken ova
15	17.8	80.9	10808	1 AE011563	AE011563 Leptospi
16	17.8	80.9	47466	8 AP002061	AP002061 Arabidops
17	17.8	80.9	65716	2 AC025143	AC025143 Homo sapi
18	17.8	80.9	110000	1 AE017300	AE017300 Leptospi
19	17.8	80.9	147837	9 AC146849	AC146849 Pan trogl

c	20	17.8	80.9	172438	9	AP005269	AP005269 Pan trogl
c	21	17.8	80.9	176109	9	AC003992	AC003992 Homo sapi
c	22	17.8	80.9	191457	2	BX469927	BX469927 Homo sapi
c	23	17.4	79.1	123570	9	AC022846	AC022846 Homo sapi
c	24	17.4	79.1	138943	8	AC133709	AC133709 Medicago
c	25	17.4	79.1	138945	8	AC144928	AC144928 Medicago
c	26	17.4	79.1	175590	2	AF307159	AF307159 Homo sapi
c	27	17.4	79.1	181761	2	AC090199	AC090199 Homo sapi
c	28	17.4	79.1	181944	9	AC022910	AC022910 Homo sapi
c	29	17.2	78.2	142	6	CQ074805	CQ074805 Sequence
c	30	17.2	78.2	142	6	CQ105711	CQ105711 Sequence
c	31	17.2	78.2	142	6	CQ144397	CQ144397 Sequence
c	32	17.2	78.2	142	6	CQ179873	CQ179873 Sequence
c	33	17.2	78.2	142	6	CQ227604	CQ227604 Sequence
c	34	17.2	78.2	142	6	CQ265760	CQ265760 Sequence
c	35	17.2	78.2	142	6	CQ302846	CQ302846 Sequence
c	36	17.2	78.2	142	6	CQ340018	CQ340018 Sequence
c	37	17.2	78.2	483	6	CQ065591	CQ065591 Sequence
c	38	17.2	78.2	483	6	CQ092557	CQ092557 Sequence
c	39	17.2	78.2	483	6	CQ113163	CQ113163 Sequence
c	40	17.2	78.2	483	6	CQ169967	CQ169967 Sequence
c	41	17.2	78.2	483	6	CQ214573	CQ214573 Sequence
c	42	17.2	78.2	483	6	CQ253167	CQ253167 Sequence
c	43	17.2	78.2	483	6	CQ290312	CQ290312 Sequence
c	44	17.2	78.2	483	6	CQ327270	CQ327270 Sequence
c	45	17.2	78.2	639	8	USC236138	AJ236138 Ustilago

ALIGNMENTS

RESULT 1	AX207103	Sequence 56 from Patent WO0153476.	22 bp	DNA	linear	PAT 30-AUG-2001
LOCUS	AX207103					
DEFINITION	Sequence 56 from Patent WO0153476.					
ACCESSION	AX207103					
VERSION	AX207103.1	GI:15394902				
KEYWORDS						
SOURCE						
ORGANISM	Pisum sativum (pea)					
REFERENCE	1	Bruce, W.B. and Niu, X.				
AUTHORS		Novel plant promoters and methods of use				
TITLE		Patent: WO 0153476-A 56 26-JUL-2001;				
JOURNAL		PIONEER HI-BRED INTERNATIONAL, INC. (US)				
FEATURES		Location/Qualifiers				
source	1..22	/organism="Pisum sativum"				
		/mol_type="unassigned DNA"				
		/db_xref="taxon:3888"				
ORIGIN						
Query Match	100.0%	Score 22; DB 6; Length 22;				
Best Local Similarity	100.0%;	Pred. No. 1.2;				
Matches	22; Conservative	0; Mismatches	0; Indels	0; Gaps	0;	
QY	1	GACACGTAGATGAGTCATCAC	22			
Db	1	GACACGTAGATGAGTCATCAC	22			
LOCUS	AX358110	Sequence 5 from Patent WO0194394.	22 bp	DNA	linear	PAT 13-FEB-2002
DEFINITION	Sequence 5 from Patent WO0194394.					
ACCESSION	AX358110					
VERSION	AX358110.1	GI:18674857				
KEYWORDS						
SOURCE		Agrobacterium tumefaciens (Rhizobium radiobacter)				

```

ORGANISM      Agrobacterium tumefaciens
REFERENCE      Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
AUTHORS        Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
TITLE          Jilka,J.M., Hood,E.E. and Howard,J.A.
JOURNAL        Novel plant promoter sequences and methods of use for same
                Patent: WO 0194394-A 5 13-DEC-2001;
                Prodigene, Inc. (US)
FEATURES
  source        1..22
                /organism="Agrobacterium tumefaciens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:358"

ORIGIN
Query Match      100.0%; Score 22; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GACACGTAGATGATCATCAC 22
        |||
        1 GACACGTAGATGATCATCAC 22

Db
1 GACACGTAGATGATCATCAC 22

RESULT 3
AY192160      827 bp      DNA      linear      SYN 26-FEB-2003
LOCUS         Synthetic construct multi-copy enhancer promoter element, complete
ACCESSION     AY192160
VERSION       AY192160
KEYWORDS      AY192160.1 GI:28569528
SOURCE        synthetic construct
ORGANISM      synthetic construct
REFERENCE      other sequences; artificial sequences.
AUTHORS        1 (bases 1 to 827)
TITLE          Yao,Q., Peng,R. and Xiong,A.
JOURNAL        An artificial promoter containing many copies of enhancers can
                increase phytase expression in transgenic rape
                Unpublished
REFERENCE      2 (bases 1 to 827)
AUTHORS        Yao,Q., Peng,R. and Xiong,A.
TITLE          Direct Submission
JOURNAL        Submitted (07-DEC-2002) Biotechnology Research Center, Shanghai
                Academy of Agriculture Science, Baidi Road 2901, Shanghai, Shanghai
                201106, China
FEATURES
  source        Location/Qualifiers
                1..827
                /organism="synthetic construct"
                /mol_type="genomic DNA"
                /db_xref="taxon:32630"
                1..827
                /note="multi-copy enhancer promoter element"
                1..90
                /note="contains 3 copies of 30bp element from glycine
                promoter"
                91..160
                /note="contains 3 copies of 22bp W1 seed-specific
                cis-element from pea lectin promoter"
                161..376
                /note="contains 3 copies of 68bp UNS from phaseolin
                promoter"
                501..827
                /note="contains CaMV mini promoter (-60) and TMV omega
                enhancer"

ORIGIN
Query Match      100.0%; Score 22; DB 12; Length 827;
Best Local Similarity 100.0%; Pred. No. 0.94;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GACACGTAGATGATCATCAC 22
        |||
        1 GACACGTAGATGATCATCAC 22

Db
1 GACACGTAGATGATCATCAC 22

RESULT 4
PEALBEB      1659 bp      mRNA      linear      PLN 27-APR-1993
LOCUS         Pea PEB1 gene encoding lectin, complete cds.
ACCESSION     M18160
VERSION       M18160.1 GI:169112
KEYWORDS      lectin.
SOURCE        Pisum sativum (pea)
ORGANISM      Pisum sativum
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae;
                Pisum.
REFERENCE      1 (bases 1 to 1659)
AUTHORS        Kaminski,P.A., Buffard,D. and Stroberg,A.D.
TITLE          The pea lectin gene family contains only one functional gene
JOURNAL        Plant Mol. Biol. 9, 497-507 (1987)
COMMENT        Original source text: Pea seed, cDNA to mRNA, clone lambda-I-101.
FEATURES
  source        Location/Qualifiers
                1..1659
                /organism="Pisum sativum"
                /mol_type="mRNA"
                /db_xref="taxon:3888"
                <1..1659
                /product="lectin mRNA"
                463..1290
                /note="lectin"
                /codon_start=1
                /protein_id="AAA33676.1"
                /db_xref="GI:169113"
                /translation="MASLQTMISFYALFLSLTLTLFFKVNSTETTSFLITSPD
                QGNLFQGDGYTKKTLTKAVKTVGVRAVSSPIHMDRETVGVANFTSFTFVIN
                AENSYNVADGFTFFIAPVDTRKQNGGYGVNSAEVDKQTQVAVEDDFYNAADP
                SNRDHIGIDVNSIKSVNTKSMKMLONGEAVVIAFNATVTLVSLTPNSLSEENV
                TSYLSDVVSXKDVPEWVRIGFATTTABEYAAHVLVSNSTHSELSTSSKQADA"

ORIGIN
Query Match      100.0%; Score 22; DB 8; Length 1659;
Best Local Similarity 100.0%; Pred. No. 0.89;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GACACGTAGATGATCATCAC 22
        |||
        1 GACACGTAGATGATCATCAC 413

Db
392 GACACGTAGATGATCATCAC 413

RESULT 5
A14134      1701 bp      DNA      linear      PAT 15-JAN-2003
LOCUS         promoter and lectin gene.
ACCESSION     A14134
VERSION       A14134.1 GI:490034
KEYWORDS
SOURCE        Pisum sativum (pea)
ORGANISM      Pisum sativum
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae;
                Pisum.
REFERENCE      1 (bases 1 to 1701)
AUTHORS        Hepher,A., Edwards,G.A. and Gatehouse,J.A.
TITLE          Improvements relating to transgenic plants
JOURNAL        Patent: EP 0351924-A 1 24-JAN-1990;
                SHELF INTERNATIONAL RESEARCH MATSCHAPPIJ B.V.; Nickerson Seeds
                Limited
FEATURES
  source        Location/Qualifiers
                1..1701
                /organism="Pisum sativum"
                /mol_type="unassigned DNA"
                /db_xref="taxon:3888"

```

gene 473..1300
/gene="leca"
473..1300
/gene="leca"
/codon_start=1
/product="lectin"
/protein_id="CAA01149.1"
/db_xref="GI:490035"
/translation="MASHOTOMISFYALFSLITLTITLFPKNSGETTSPLITKPSD
OONLFGQDGTTEKTLTKKXKNTVGKALYSPIHIMDRGTGVNAFVTSFTVIN
APNSYNVADGFTFTPLAPVDTKPTQGTGVLGVNSAEYDKTQTVNAFVTSFTVIN
SNRPHIGIDVNSISVNTKSKLONGBEAVNVIAFNATVTLTSLTYPSLSEENV
TSYLSDVVSLKDVPEWVRIGFSATGAEVAAHEVLSMFSHSELSTSSSKOADA"

Query Match 100.0%; Score 22; DB 6; Length 1701;
Best Local Similarity 100.0%; Pred. No. 0.89;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GACAGTAGAATGAGTCATCAC 22
|||
414 GACAGTAGAATGAGTCATCAC 435

RESULT 6
LOCUS PSIECA 1701 bp DNA linear PLN 31-MAR-1995
DEFINITION Pisum sativum leca gene for lectin.
ACCESSION Y00440
VERSION Y00440.1 GI:20769
KEYWORDS leca gene; lectin.
SOURCE Pisum sativum (pea)
ORGANISM Pisum sativum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae;
Pisum.
1 (bases 1 to 1701)
Gatchouse, J.A., Bown, D., Evans, I.M., Gatchouse, L.N., Jones, D.,
Preston, P., and Croy, R.R.
Sequence of the seed lectin gene from pea (Pisum sativum L.)
JOURNAL Nucleic Acids Res. 15 (18), 7642 (1987)
MEDLINE 88015625
PUBMED 3658708
2 (bases 1 to 1701)
Gatchouse, J.A.
Direct Submission
Submitted (07-SEP-1987) Gatchouse J.A., Botany Department,
University of Durham, South Road, Durham, DH1 3LE, England
FEATURES
source
1..1701
/organism="Pisum sativum"
/mol_type="genomic DNA"
/strain="Peldham First"
/db_xref="taxon:3888"
/tissue_type="seed"
408..413
/note="TATA-box"
473..1300
/codon_start=1
/product="lectin-precursor (AA -30 to 245)"
/protein_id="CAA68497.1"
/db_xref="GI:20770"
/db_xref="GOA:P02867"
/db_xref="UniProt/Swiss-Prot:P02867"
/translation="MASHOTOMISFYALFSLITLTITLFPKNSGETTSPLITKPSD
OONLFGQDGTTEKTLTKKXKNTVGKALYSPIHIMDRGTGVNAFVTSFTVIN
APNSYNVADGFTFTPLAPVDTKPTQGTGVLGVNSAEYDKTQTVNAFVTSFTVIN
SNRPHIGIDVNSISVNTKSKLONGBEAVNVIAFNATVTLTSLTYPSLSEENV
TSYLSDVVSLKDVPEWVRIGFSATGAEVAAHEVLSMFSHSELSTSSSKOADA"
sig_peptide
473..562
/note="signal peptide (AA -30 to -1)"
mat_peptide
563..1297

misc_feature
1359..1365
/note="polyA signal"
polyA_site
1422
/note="pot. polyA site"
polyA_site
1446
/note="alt. pot. polyA site"

Query Match 100.0%; Score 22; DB 8; Length 1701;
Best Local Similarity 100.0%; Pred. No. 0.89;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GACAGTAGAATGAGTCATCAC 22
|||
414 GACAGTAGAATGAGTCATCAC 435

RESULT 7
LOCUS PSIECA 3360 bp DNA linear PLN 31-OCT-1994
DEFINITION P.sativum psl gene for psl lectin.
ACCESSION X66368
VERSION X66368.1 GI:562782
KEYWORDS lectin; psl gene; Psl lectin.
SOURCE Pisum sativum (pea)
ORGANISM Pisum sativum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae;
Pisum.
1 (bases 1 to 3360)
de Pater, S., Pham, K., Chua, N.H., Memelink, J., and Kijne, J.
A 22-bp fragment of the pea lectin promoter containing essential
TGAC-like motifs confers seed-specific gene expression
Plant Cell 5 (8), 877-886 (1993)
MEDLINE 94004020
PUBMED 8400870
2 (bases 1 to 1713)
Hoedemaecker, F.J., Richardson, M., Diaz, C.L., de Pater, B.S. and
Kijne, J.W.
Pea (Pisum sativum L.) seed isolectins 1 and 2 and pea root lectin
result from carboxypeptidase-like processing of a single gene
product
Plant Mol. Biol. 24 (1), 75-81 (1994)
MEDLINE 94154245
PUBMED 8111028
3 (bases 1 to 3360)
de Pater, S., Katsagiri, F., Kijne, J., and Chua, N.H.
bzlp proteins bind to a palindromic sequence without an ACGT core
located in a seed-specific element of the pea lectin promoter
Plant J. 6 (2), 133-140 (1994)
4
de Pater, B.S.
Direct Submission
Submitted (19-MAY-1992) B.S. de Pater, Center for Phytotechnology,
Botanical Laboratory, Nonnensteeg 3, 2311 VJ Leiden, THE
NETHERLANDS
revised by [3] MAT
5 (bases 1 to 3360)
de Pater, B.S.
Direct Submission
Submitted (27-OCT-1994) B.S. de Pater, Center for Phytotechnology,
Cultus Laboratorium, Wassenaarseweg 64, 2333 AL Leiden, THE
NETHERLANDS
On Nov 1, 1994 this sequence version replaced gi:20803.
location/Qualifiers
1..3360
/organism="Pisum sativum"
/mol_type="genomic DNA"
/variety="Peldham First"


```

repeat_region complement(10614..11362)
/rpt_family="LIMC4a"
repeat_region 11364..11700
/rpt_family="L2"
repeat_region 11903..12038
/rpt_family="L2"
repeat_region 12793..12820
/rpt_family="AT_rich"
repeat_region complement(13170..13201)
/rpt_family="U2"
repeat_region complement(13287..13340)
/rpt_family="MERSA"
repeat_region 13395..13443
/rpt_family="MERSA"
repeat_region complement(13401..13442)
/rpt_family="MERSB"
repeat_region 14965..15007
/rpt_family="(TCTA)n"
repeat_region complement(15008..15294)
/rpt_family="AluSx"
repeat_region 15296..15317
/rpt_family="AT_rich"
repeat_region 15572..15759
/rpt_family="MERSA"
repeat_region complement(16050..16156)
/rpt_family="LTR1"
repeat_region complement(16219..16611)
/rpt_family="HUBS-P2"
repeat_region complement(16612..17391)
/rpt_family="LTR1"
repeat_region 17546..18358
/rpt_family="HAL1"
repeat_region 18486..18647
/rpt_family="MERSA"
repeat_region complement(18654..19102)
/rpt_family="MLT1C"
repeat_region 19131..19273
/rpt_family="HAL1"
repeat_region 19554..19621
/rpt_family="HAL1"
repeat_region 19818..19845
/rpt_family="AT_rich"
repeat_region complement(20165..20627)
/rpt_family="LIM4"
repeat_region complement(20628..21367)
/rpt_family="LIPAT"
repeat_region complement(21368..21519)
/rpt_family="LIM4"
repeat_region complement(21872..22228)
/rpt_family="MERS3"
repeat_region complement(22232..22327)
/rpt_family="LTR4B"
repeat_region complement(22378..22857)
/rpt_family="MERS3"
repeat_region 24790..24827
/rpt_family="(T)n"
repeat_region 25131..25201
/rpt_family="(TATATG)n"
repeat_region complement(25277..26219)
/rpt_family="LIM4"
repeat_region complement(26222..26375)
/rpt_family="LTR9"
repeat_region complement(26404..26909)
/rpt_family="LTR9"
repeat_region complement(26911..34208)
/rpt_family="HUBS-P3b"
repeat_region 34209..34324
/rpt_family="(CATTC)n"
repeat_region complement(34325..34399)
/rpt_family="HUBS-P3b"
repeat_region complement(34429..34947)
/rpt_family="LTR9"
repeat_region complement(34952..37378)

```

```

repeat_region /rpt_family="LIM4"
complement(37433..38509)
repeat_region /rpt_family="LIM4"
complement(38536..39436)
repeat_region /rpt_family="LIM4"
complement(39832..39969)
repeat_region /rpt_family="FLAM C"
complement(40224..40514)
repeat_region /rpt_family="AluSg"
complement(40897..41020)
repeat_region /rpt_family="MIR"
41370..41508
repeat_region /rpt_family="L2"
41513..41554
repeat_region /rpt_family="AT_rich"
complement(42256..42369)
repeat_region /rpt_family="MLT1H"

Query Match 85.5%; Score 18.8; DB 9; Length 136492;
Best Local Similarity 90.9%; Pred. No. 36;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GACACGTAGATGAGTCATCAC 22
Db 136431 GACACGTAGATGAGTCATCAC 136452

RESULT 9
AC090660 195156 bp DNA linear PRI 28-NOV-2001
LOCUS Homo sapiens chromosome 18, clone RP11-86901, complete sequence.
DEFINITION AC090660
AC090660.3 GI:17047149
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Graniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 195156)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Unpublished
2 (bases 1 to 195156)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
Barra, N., Baetien, V., Boguslavsky, L., Boukhalter, B., Brown, A.,
Camata, V., Campopiano, A., Choepel, Y., Colangelo, M., Collins, S.,
Collimore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Faro, S., Ferreira, P., Fitzhugh, M., Gage, D., Galegan, J.,
Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Karatas, A., Larocque, K., Lamazeres, R., Landers, T.,
Lehoczky, V., Levine, R., Liu, G., MacLean, C., MacDonald, P.,
Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,
McPheeters, R., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R.,
Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M.,
Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P.,
Sounez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Travers, M., Travis, N., Trigilio, J., Vasilev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zainoun, J.,
Zemek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (07-MAR-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 195156)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barra, N., Baetien, V., Boguslavsky, L., Boukhalter, B.,
Brown, A., Camata, V., Campopiano, A., Choepel, Y., Colangelo, M.,
Choepel, Y., Colangelo, M., Collins, S., Collimore, A., Cooke, P.,
Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,

```

[illegible]

```

repeat_region      31484..31525
                    /rpt_family="AT_rich"
repeat_region      complement(32227..32340)
                    /rpt_family="MULTIH"
repeat_region      complement(32416..32685)
                    /rpt_family="MULTIH"
repeat_region      complement(32765..32914)
                    /rpt_family="MIR"
repeat_region      33798..33825
                    /rpt_family="(TA)n"

```

```

Query Match      85.5%; Score 18.8; DB 9; Length 195156;
Best Local Similarity 90.9%; Pred. No. 36;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Gy 1 GACACGTAGATGATCATCATC 22
    |||||
Db 126402 GACACGTAGATGATCATCATC 126423

```

```

RESULT 10
AP001593/c      198127 bp   DNA      linear   HTG 26-JUL-2000
LOCUS
DEFINITION      Homo sapiens chromosome 18 clone RP11-86901 map 18q21, WORKING
DRAT SEQUENCE, 11 unordered pieces.
ACCESSION      AP001593
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 198127)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.,
Homo sapiens 198,127 genomic DNA of 18q21
Published Only in Database (2000)
2 (bases 1 to 198127)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.,
Direct Submission
Submitted (31-MAR-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-15-1 Kitesato, Sagamihara, Kanagawa 228-8555, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-42-778-9923, Fax:81-42-778-9924)
On Jul 26, 2000 this sequence version replaced gi:8117417.
-----
Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
-----
Project Information
Center project name: HumDrafit18
Center clone name: RP11-86901
-----
Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 195298 bases at least Q40
Consensus quality: 196428 bases at least Q30
Consensus quality: 196782 bases at least Q20
Insert size: 197127; sum-of-ctnigs
Quality coverage: 9.52x in Q20 bases; sum-of-ctnigs

```

NOTE: This is a 'working draft' sequence. It currently consists of 11 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1 58065 contig of 58065 bp in length

```

58166 92043 contig of 33878 bp in length
92144 111587 contig of 19444 bp in length
111688 129552 contig of 17865 bp in length
129653 149317 contig of 19665 bp in length
149418 164242 contig of 14825 bp in length
164343 175076 contig of 10734 bp in length
175177 185856 contig of 10680 bp in length
185957 195249 contig of 9233 bp in length
195350 196868 contig of 1519 bp in length
196969 198127 contig of 1159 bp in length
198127 198127 contig of 1159 bp in length.
Location/Qualifiers
1..198127
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="18"
/map="18q21"
/clone="RP11-86901"
1..58065
/note="assembly_fragment"
58166..92043
/note="assembly_fragment"
92144..111587
/note="assembly_fragment"
111688..129552
/note="assembly_fragment"
129653..149317
/note="assembly_fragment"
149418..164242
/note="assembly_fragment clone_end:SP6 vector_side:left"
164343..175076
/note="assembly_fragment clone_end:T7 vector_side:right"
175177..185856
/note="assembly_fragment"
185957..195249
/note="assembly_fragment"
195350..196868
/note="assembly_fragment"
196969..198127
/note="assembly_fragment"

```

```

ORIGIN
Query Match      85.5%; Score 18.8; DB 2; Length 198127;
Best Local Similarity 90.9%; Pred. No. 35;

```

	Matches	20;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0
Qy	1	GACACTAGAAATGATCATTCCAC	22							
Db	179594	GACACGTAGAATTAGTACCAC	179573							
RESULT 11										
LOCUS	CQ870293		349980 bp	DNA						
DEFINITION	Sequence 714 from Patent WO2004074320.									
ACCESSION	CQ870293									
VERSION	CQ870293.1		GI:51999890							
KEYWORDS										
SOURCE										
ORGANISM	Homo sapiens (human)									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.									
TITLE	1 Morris,D.W., Morris,D.W. and Malandro,M.S.									
JOURNAL	Novel therapeutic targets in cancer									
FEATURES	Patent: WO 2004074320-A 714 02-SEP-2004;									
SOURCE	Sagres Discovery, Inc. (US)									
	Location/Qualifiers									
	1..349980									
	/organism="Homo sapiens"									
	/mol_type="unassigned DNA"									
	/db_xref="taxon:9606"									
	/note="seq 714, original 779.603 bases, replaced by-seq									
	714, from 0.000.001 to 0.349.980-seg 905, from 0.300.001									
	to 0.649.980-seg 906, from 0.600.001 to 0.779.603"									
ORIGIN										
Query Match	85.5%; Score 18.8; DB 6; Length 349980;									
Beet Local Similarity	90.9%; Pred.No. 34;									
Matches	20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;									
Qy	1	GACACGTAGATGATCATTCAC	22							
Db	142953	GACACGTAGAATTAGTACCAC	142974							
RESULT 12										
LOCUS	ACI25804/C									
DEFINITION	Rattus norvegicus clone CH230-818, WORKING DRAFT SEQUENCE, 3									
ACCESSION	ACI25804		226810 bp	DNA						
VERSION	ACI25804.4		GI:30467256							
KEYWORDS	HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.									
SOURCE	Rattus norvegicus (Norway rat)									
ORGANISM	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;									
	Rattus.									
	1 (bases 1 to 226810)									
REFERENCE	Muzny,D.Marie, Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,									
AUTHORS	Allen,C., Allen,H., Alsbrook,S., Amin,A., Anguiano,D,									
	Anyalabechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden H,									
	Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F,									
	Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,									
	Bryant,N., Buhey,C., Burch,P., Burrell,K., Calderon,E.,									
	Cardenas,V., Carter,K., Cavazos,I., Cassar,H., Center,A.,									
	Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,									
	Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L,									
	Davila,M.O., Davis,C., Davy-Carroli,L., De Anda,C., Dedertich,D.,									
	Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,									
	Draper,H., Dugan-Rocha,S., Dunn,C., Durish,K., Duval,B., Evans,K.,									
	Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,									
	Fernandez,S., Finley,M., Flaggs,N., Forbes,L., Foster,M., Foster,P,									
	Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,									
	Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,M.,									
	Gunnarete,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,									

Harvey, Y., Havlik, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladik, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hu, Y. S., Hume, J., Idelir, D., Jackson, A., Jackson, J., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolyet, A., Kapachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kow, C., Kraft, C. L., Ledow, H., Levay, J., Lewis, L., Li, Z., Liu, J., Liu, F., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshaw, L., Loulseg, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Manthey, S., McLeod, M. P., McNeill, T. Z., Meenan, E., Milosavljevic, A., Miner, G., Ming, E., Montanoy, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundaas, M., Murphy, M., Nair, L., Narkevicius, C., Neal, D., Newton, N., Nguyen, N., Norrie, A., Nwokoleh, O., Okwuonu, G., Olarnpungsoong, A., Pal, S., Parke, K., Paternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Reiger, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, C., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Smaj, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uemami, K., Valas, R., Vera, V., Villanias, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willison, R., Wleczek, R., Wooden, H., Woley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G., and Gibbs, R. A.

Direct Submission

Unpublished

2 (bases 1 to 226810)

Worley, K. C.

Direct Submission

Submitted (02-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 226810)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 9, 2003 this sequence version replaced gi:24940821.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separate sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GCTT

Center clone name: CH230-818

----- Summary Statistics

Assembly program: Atlas

Consensus quality: 218425 bases at least Q40

Consensus quality: 220364 bases at least Q30

Consensus quality: 221808 bases at least Q20

Estimated insert size: 226184; sum-of-contigs estimation

Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.ngsc.bcm.tmc.edu/docs/genbank_drafc_data.html).
NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 224161: contig of 224161 bp in length
* 224162 224261: gap of unknown length
* 224262 225628: contig of 1367 bp in length
* 225629 225728: gap of unknown length
* 225729 226810: contig of 1082 bp in length.

FEATURES
source
1. 226810
Location/Qualifiers
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-818"

misc_feature
1. 1284
/note="wgs contig"
misc_feature
223161..224161
/note="wgs_contig"

ORIGIN

Query Match 83.6%; Score 18.4; DB 2; Length 226810;
Best Local Similarity 95.0%; Pred. No. 58;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ACACGTAGATGATCATCA 21
Db 9048 ACACGTAGATGATCATCA 9029

RESULT 13
BX927258 146407 bp DNA linear VRT 27-MAR-2004
LOCUS zebrafish DNA sequence from clone CH211-279M15 in linkage group 20,
DEFINITION complete sequence.
ACCESSION BX927258 GI:45774169
VERSION BX927258.8 GI:45774169
KEYWORDS HTG.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 146407)
Babbage,S.
Direct Submision
Submitted (27-MAR-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zf1sh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Mar 28, 2004 this sequence version replaced gi:45535664.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zf1sh-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhifeng Bao and Sean Eddy, submitted), and those beginning 'dir' were identified by Rick Waterman (Stephen Johnson lab, Washu). For further information see http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml CH211-279M15 is from a CHORI-211 BAC library

FEATURES

source
1. 146407
Location/Qualifiers
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-279M15"
/clone_lib="CHORI-211"

ORIGIN

Query Match 81.8%; Score 18; DB 5; Length 146407;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACGTAGATGATCA 18
Db 7456 GACACGTAGATGATCA 7439

RESULT 14
CHROVAL5A/C 5322 bp DNA linear VRT 28-APR-1993
LOCUS Chicken ovalbumin gene, 5' flank.
DEFINITION M29020
ACCESSION M29020 GI:212507
VERSION M29020.1 GI:212507
KEYWORDS ovalbumin.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 5322)
Kaye,J.S.
Unpublished (1989)
Original source text: Chicken DNA, clones pBR EcoRI[6,7].
Draft entry and computer-readable sequence for [1] kindly submitted by J.S.Kaye, 14-OCT-1989.

Location/Qualifiers
1. 5322
/organism="Gallus gallus"
/mol_type="genomic DNA"
/db_xref="taxon:9031"

FEATURES

source
1. 5322
Location/Qualifiers
/organism="Gallus gallus"
/mol_type="genomic DNA"
/db_xref="taxon:9031"

ORIGIN

Query Match 80.9%; Score 17.8; DB 5; Length 5322;
Best Local Similarity 90.5%; Pred. No. 1.6e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GACACGTAGATGATCA 21
Db 1539 GACACGTAGATGATCA 1519

```
RESULT 15
AB011563
LOCUS
DEFINITION
AB011563 Leptospira interrogans serovar lai str. 56601 chromosome I, section
372 of 397 of the complete sequence.
AB011563
VERSION
AB011563.1 GI:24198148
SOURCE
ORGANISM
Leptospira interrogans serovar lai str. 56601
Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
REFERENCE
1 (bases 1 to 10808)
Ren, S.X., Fu, G., Jiang, X.G., Zeng, R., Miao, Y.G., Xu, H., Zhang, Y.X.,
Xiong, H., Lu, G., Lu, J.F., Jiang, H.Q., Jia, J., Tu, Y.F., Jiang, J.X.,
Gu, W.Y., Zhang, Y.Q., Cai, Z., Sheng, H.H., Yin, H.F., Zhang, Y.,
Zhu, G.F., Wan, M., Huang, H.L., Qian, Z., Wang, S.Y., Ma, W., Yao, Z.J.,
Shen, Y., Qiang, B.Q., Xia, Q.C., Guo, X.K., Danchin, A., Saint
Giron, I., Somerville, R.L., Wen, Y.M., Shi, M.H., Chen, Z., Xu, J.G.
and Zhao, G.P.
Unique physiological and pathogenic features of Leptospira
interrogans revealed by whole-genome sequencing
Nature 422 (6934), 888-893 (2003)
12712204
2 (bases 1 to 10808)
Ren, S., Fu, G., Jiang, X., Zeng, R., Xiong, H., Lu, J., Lu, G., Jiang, H.,
Ding, Y., Jia, J., Tu, Y., Gu, W., Cai, Z., Sheng, H., Yin, H., Zhang, Y.,
Zhu, G., Wang, S., Shen, Y., Qiang, B., Chen, Z., Wen, Y., Xu, J. and
Zhao, G.
Direct Submission
Submitted (12-MAR-2002) Chinese National Human Genome Center at
Shanghai, 250 Bi Bo Road, Shanghai 201203, China
Updated information will be available at our World Wide Web site
(http://www.chgc.sh.cn/lep/). Comments to the authors are
appreciated.
FEATURES
source
1. 10808
/organism="Leptospira interrogans serovar lai str. 56601"
/mol_type="genomic DNA"
/strain="56601"
/serovar="lai"
/db_xref="taxon:189518"
/chromosome="I"
/notes="Strain 56601 is maintained by the Institute of
Epidemiology and Microbiology, Chinese Academy of
Preventive Medicine"
228..722
/locus_tag="LA4087"
228..722
/locus_tag="LA4087"
/codon_start=1
/product="hypothetical protein"
/transl_table=11
/protein_id="AAN51285.1"
/db_xref="GI:24198149"
/translacion="MFLFRLLEKKKKLLIILFFSSCKKDPDPAVTEQEKDILIQ
ILMENESIHFKLEKEKIPDNTOLIAHVEVLSLNGKLSQSEKQNSLKDIEDVD
EKFQGVSSFSSENLASIKLAGGTGPNRKYCPMVKNTWQSTKIRNPADPMDCG
DIYH"
701..1690
/locus_tag="LA4088"
701..1690
/locus_tag="LA4088"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAN51286.1"
/db_xref="GI:24198150"
/translacion="MMGPSTLNVETCLDEIKTRCKIGVSLGEPGRANCROVLTNEN
SSIRSCJAGEKMAFAFGKAGINVASAIFALPVTLEERISKGRVYDELITLIFS
IIEPEKKEVIVFSNKLKYLIIIVYNYQOTEDBRKSLKRNASFLDIREAVAGS
SBERGCFIAYCKEAKKDEPERASQFTITVPSAVSDLRNLAGLNVEEKELYLKDG
IYELPQIPIGIRHISLCPEDDEPIFLIISTWEELVLRKOQIIESSHAILEKYSKGL
NMSLPGDLSVLELSMEILGIFFEKELIGRSEKRLIKELLFKHKHKNELT"
gene
CDS
complement(1699..2523)
/locus_tag="LA4089"
complement(1699..2523)
/locus_tag="LA4089"
/codon_start=1
/transl_table=11
/product="phenazine biosynthesis protein, PhzF family"
/protein_id="AAN51287.1"
/db_xref="GI:24198151"
/translacion="MKSIIQWBCSKAIFMFKLQIDSPAEKVFQGNPAVCPDQEWI
SDMMOKIASENNLSERAPVVGESRRIRWFTPLREVDLCGATLAAVYHLFRGEV
IGKLNQGSISGLNFKENITLYLPSPKPEKVTPEILNFSISKEVFKSDY
MLATYENQJLKLEYSVEGRLENTLGIATCGNDSYDLSRFPAPSGLVEDPYTG
SAACTLIPIFWSERLGGKIMKAYOTSARGGLFCEDIGSERVIRGTCVPIADGIEI"
2752..2868
/locus_tag="LA4090"
2752..2868
/locus_tag="LA4090"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAN51288.1"
/db_xref="GI:24198152"
/translacion="MKKFRVYVPTLIEFVHKIPGMSALPQKPLFRNQFTII"
complement(3535..5007)
/locus_tag="LA4092"
complement(3535..5007)
/gene="purP"
/locus_tag="LA4092"
/BC_number="2.4.2.14"
/notes="Glutamine phosphoribosylpyrophosphate
amidotransferase"
/codon_start=1
/transl_table=11
/product="Amidophosphoribosyltransferase"
/protein_id="AAN51289.1"
/db_xref="GI:24198153"
/translacion="MSRIIDKSRIRROQDDKPEKECAIFGIFNSSEANSFTYGLYS
NQHROQSSGIVSSDGEHLRYAGMGVAAHFTETKLELQGNALIGNRSTGASG
LMAQPLRVESHGAVSLAHNGNLVNSWELSKQEKESIQTIDSEVIVHMAKSG
ETDFLSALSSALKVKGAVSLVILTKTOLIVRDNGRPLVMGRREDGIVPASSTC
AFDIDTKYERDVERGEMIVVDKGNSTYVFPASBGLCFEYIYPARDPSSIFGES
VYKVRKNGRLIARELPVPADVVPYPSANIALGVAESGISYSGILNSHIGRT
PIEPQKIRDDGAKTKVNVKRVVGVVVDISIMGTSSKRIKQIRNAKKEIH
LAVSAPPTISPEYIGDILPTNELIAAHTHEELRKYLVDISAYLSVESMRAVMDH
KGGGFCNACFTAOYVPEFQSELGNQSLFKEYQVEERVVY"
4950..6215
/locus_tag="LA4091"
4950..6215
/locus_tag="LA4091"
/codon_start=1
/transl_table=11
/product="Ribonuclease D"
/protein_id="AAN51290.1"
/db_xref="GI:24198154"
/translacion="MVCHPGILVPLFWIYREBDSASFLSKQVAFPMQINSYIVVD
TARSLQVLTINLSQADSISVDTESSGYTTFKVCGLQISAKGNVITIDPLKQNES
LGNLPEDKILIKTIFSAIDIDIKALKKQFGQFOQIADTGSRLDHEQYSTLYVDY
YHKIKLSKQSGNSWNERPLEKQLOQAALDTYVLETIWEKMEELIKNLVYEAISE
FEKIASSEPGSENSIMDKFPEILYESADERRFIVDTLVFRDQSKLKNAPFVLE
NEKVVLLMKARMDKKLTELIVGKDAETLQIYANBGPPIQSELPKKGELTYNEE
GERFKRLIRIRETIMSIRMSHOMBPENKMLIALSQNPRTLDLKKMLFSEKVIH
YPSILALANVPRESNIKILIPINKFE"
6247..6918
/locus_tag="LA4093"
6247..6918
/locus_tag="LA4093"
/codon_start=1
/transl_table=11
/product="MutT/nudix family protein"
/protein_id="AAN51291.1"
/db_xref="GI:24198155"
/
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2005, 23:35:01 ; Search time 2081.3 Seconds
(without alignments)
548.661 Million cell updates/sec

Title: US-10-086-062-4

Perfect score: 30
Sequence: 1 ctggaccctctcgactcgagagcttcgct 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:
1: gb_ests1:
3: gb_ests2:
4: gb_ests3:
5: gb_ests4:
6: gb_ests5:
7: gb_ests6:
8: gb_ests7:
9: gb_ests8:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20.6	68.7	1091	5	BM905373 AGENCOURT
2	20.4	68.0	804	7	BG430455 602502021
3	20.2	67.3	877	7	CK873283 AGENCOURT
4	20.2	67.3	1218	4	BG474756 602517583
5	20	66.7	838	7	CN127674 RHOMI 24
6	19.6	65.3	634	9	CE429361 CIGR-986-
7	19.4	64.7	123	2	BF869056 MR4-TN011
8	19.4	64.7	168	9	BF869056 MR4-TN011
9	19.4	64.7	240	4	BF869056 MR4-TN011
10	19.4	64.7	245	2	BF869056 MR4-TN011
11	19.4	64.7	254	2	BF869056 MR4-TN011
12	19.4	64.7	469	4	BF869056 MR4-TN011
13	19.4	64.7	511	4	BF869056 MR4-TN011
14	19.4	64.7	541	4	BF869056 MR4-TN011
15	19.4	64.7	564	4	BF869056 MR4-TN011
16	19.4	64.7	588	1	BF869056 MR4-TN011
17	19.4	64.7	688	1	BF869056 MR4-TN011
18	19.4	64.7	893	7	BF869056 MR4-TN011
19	19.4	64.7	893	7	BF869056 MR4-TN011
20	19	63.3	394	5	BY018367 BY018367
21	19	63.3	453	5	BY018367 BY018367
22	19	63.3	457	6	BY018367 BY018367
23	19	63.3	467	8	BY018367 BY018367
24	19	63.3	498	5	BY018367 BY018367

25	19	63.3	504	5	BX087591 BX087591
26	19	63.3	517	7	CK518339 CK518339
27	19	63.3	522	8	AO229833 HS 2027 A
28	19	63.3	525	7	CK885179 SGP165556
29	19	63.3	526	2	BE684663 BE684663
30	19	63.3	535	5	BX301244 BX301244
31	19	63.3	544	7	CK898692 SGP162270
32	19	63.3	552	5	BX317355 BX317355
33	19	63.3	558	5	BX888009 BX888009
34	19	63.3	586	6	AZ806755 AZ806755
35	19	63.3	598	8	CA343648 CA343648
36	19	63.3	625	4	BM172236 BM172236
37	19	63.3	630	5	BX877903 BX877903
38	19	63.3	633	5	BX317619 BX317619
39	19	63.3	680	5	BX319301 BX319301
40	19	63.3	709	2	AM743305 AM743305
41	19	63.3	730	5	BX861731 BX861731
42	19	63.3	751	5	BX297481 BX297481
43	19	63.3	752	6	CA353623 CA353623
44	19	63.3	756	2	BF338402 BF338402
45	19	63.3	763	5	BX873117 BX873117

ALIGNMENTS

RESULT 1	BM905373	1091 bp	mRNA	linear	EST 12-MAR-2002
LOCUS	AGENCOURT_6698806	NIH_MGC_72	Homo sapiens	CDNA clone	IMAGE:5558278
DEFINITION	5', mRNA sequence.				
ACCESSION	BM905373	GI:19355752			
VERSION	BM905373.1	GI:19355752			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cga@bbs-remail.nih.gov Tissue Procurement: ATCC/DCTD/DTF CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Plate: LLM1280 row: p column: 23 High quality sequence stop: 573. Location/Qualifiers 1..1091 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5558278" /tissue_type="melanocytic melanoma" /lab_host="DH10B (phage-resistant)" /clone_id="NIH_MGC_72" /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: Nci; Site 2: Salt; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."				
FEATURES	source				
ORIGIN					
Query Match	68.7%	Score 20.6	DB 5	Length 1091	
Best Local Similarity	85.2%	Pred. No. 3.1e+02			
Matches	23	Conservative 0	Mismatches 4	Indels 0	Gaps 0
QY	1 CTGACCCCTCTCGACTCGAGAGTTCC 27				

Db 900 CTGAGCCCTCTCGACTCGAGATCCC 926

RESULT 2
BG430455/c 804 bp mRNA 1linear EST 14-MAR-2001
LOCUS 602502021F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4615756 5',
DEFINITION mRNA sequence.
ACCESSION BG430455
VERSION BG430455.1 GI:13337063
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 804)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: CLOUTCH Laboratories, Inc.
CDNA Library Preparation: CLOUTCH Laboratories, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM1369 row: i column: 05
High quality sequence stop: 742.
Location/Qualifiers
1. 804
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4615756"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_75"
/note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site 1:
SfiI (ggccgctcggcc); Site 2: SfiI (ggccatgagcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGGCATTATGAGCC-3' and 3' adaptor sequence:
5'-ATTGAGGCGCCGAGGCGCGGAGATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.65
kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 68.0%; Score 20.4; DB 4; Length 804;
Best Local Similarity 80.0%; Pred. No. 3.7e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTGAGCCCTCTCGACTCGAGATCGGCT 30
Db 559 CTGAGCCCTCTCGACTCGAGATCTCTCT 530

RESULT 3
CK873283/c 877 bp mRNA 1linear EST 09-MAR-2004
LOCUS AGENCOURT 19436603 NIH_ZGC_16 Danio rerio cDNA clone IMAGE:7215126
DEFINITION 5', mRNA sequence.
ACCESSION CK873283
VERSION CK873283.1 GI:45289381
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 877)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs-remail.nih.gov
Tissue Procurement: Will Talbot
CDNA Library Preparation: Dr. Yutaka Suzuki and Dr. Sumio Sugano
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM5111 row: h column: 04
High quality sequence start: 13
High quality sequence stop: 685.
Location/Qualifiers
1. 877
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:79955"
/clone="IMAGE:7215126"
/issue_type="13 pooled, includes stomach, intestine,
liver and pancreas"
/lab_host="DH10B Tona"
/clone_lib="NIH_ZGC 16"
/note="Organ: gut; Vector: pME18S-FL3, Site 1: DraIII;
Site 2: DraIII; 1st strand cDNA was primed with an
oligo(dT) primer
[GGGCTGAAGACGGCTATGTCGCTTTTCTTTTCTTTT];
double-stranded cDNA was ligated to a DraIII adaptor
[GGCCUACUGG], digested and directionally cloned into
distinct DraIII sites of the pME18S-FL3. Library was size
selected for 1.0 kb with a average insert size of ~1.2kb.
Library constructed by Yutaka Suzuki (University of Tokyo
Institute of Medical Science). Custom primers recommended
for sequencing: 5' end primer 5'-GGATGTCCTTTACTCTTA-3'
and 3' end primer 5'-CGACGTCAGCTGACACCA-3'. Note: This
is a Zebrafish Gene Collection (ZGC) library"

ORIGIN

Query Match 67.3%; Score 20.2; DB 7; Length 877;
Best Local Similarity 86.0%; Pred. No. 4.5e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 TGAACCCCTCTCGACTCGAGATTC 26
Db 817 TGAACCTGTGCTCGACTCGAGATTC 793

RESULT 4
BG474756 1218 bp mRNA 1linear EST 21-MAR-2001
LOCUS 602517583F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:4649109 5',
DEFINITION mRNA sequence.
ACCESSION BG474756
VERSION BG474756.1 GI:13407020
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 1218)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM1428 row: f column: 22
High quality sequence stop: 1.
Location/Qualifiers

1..1218
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4649109"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NH_MGC_16"
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NH_MGC library."

ORIGIN

Query Match 67.3%; Score 20.2; DB 4; Length 1218;
Best Local Similarity 88.0%; Pred. No. 4.6e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 TGACCCCTCTCGACTCGAGAGTTC 26
|||||
Db 769 TGAACCTCTCTGACTCGAGAGCTC 793

RESULT 5
LOCUS CN127674/c 838 bp mRNA linear EST 01-APR-2004
DEFINITION RHOH1_24_A01.G3_A002 Acid- and alkaline-treated roots Sorghum bicolor CDNA clone RHOH1_24_A01_A002 5', mRNA sequence.
ACCESSION CN127674
VERSION CN127674.1 GI:45953877
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
REFERENCE Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 838)
Cordonnier-Pratt M.-M., Suzuki Y., Sugano S., Klein R.R., Liang C., Sun F., Sullivan R., Lim S., Eastman A. and Pratt L.H.
An EST database from Sorghum: acid- and alkaline-treated roots unpublished (2003)
Other ESTs: RHOH1_24_A01.b1_A002
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmp@prattuga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in the Human Genome Center, University of Tokyo Institute of Medical Science; plant material and RNA prepared at Texas A & M University; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.
Seq primer: Sugs (CTTCTGCTTAAGAGCTGCG).

FEATURES

source
1..838
/organism="Sorghum bicolor"
/mol_type="mRNA"

/cultivar="BTx623"
/db_xref="taxon:4558"
/clone="RHOH1_24_A01_A002"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="Acid- and alkaline-treated roots"
/note="Organ: Root; Vector: pME18S-FL3; Site_1: XhoI; Site_2: XhoI; The library was prepared from polyA+ RNA from 8-day-old roots harvested from BTx623 sorghum seedlings grown in hydroponic culture. HCl was added to a pH of 3.0 to some seedlings, KOH to a pH of 9.0 for others. Roots were harvested 3, 12 and 27 hr after addition and pooled for RNA preparation. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pME18S-FL3 vector (5'-prime DraIII site is CACTGTGG, 3'-prime DraIII site is CACCATGG). XhoI excises the cDNA insert."

ORIGIN

Query Match 66.7%; Score 20; DB 7; Length 838;
Best Local Similarity 82.1%; Pred. No. 5.5e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CTGACCCCTCTCGACTCGAGAGTTCG 28
|||||
Db 526 CTGACCCCAAGCATCGAGAGTTCG 499

RESULT 6
LOCUS CE429361/c 634 bp DNA linear GSS 27-SEP-2003
DEFINITION tigr-gss-dog-17000362924386 Dog library Canis familiaris genomic, genomic survey sequence.
ACCESSION CE429361
VERSION CE429361.1 GI:36701435
KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 634)
Kirkness E.F., Bafna V., Halpern A.L., Levy S., Remington K., Ruesch D.B., Delcher A.L., Pop M., Wang W., Fraser C.M. and Venter J.C.
The dog genome: survey sequencing and comparative analysis Science 301 (5641), 1898-1903 (2003)

JOURNAL MEDLINE 22875432
PUBMED 14512627
COMMENT Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.

FEATURES

source
1..634
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site_1: BclXI; Libraries were prepared from peripheral blood"

ORIGIN

Query Match 65.3%; Score 19.6; DB 9; Length 634;
Best Local Similarity 84.6%; Pred. No. 8.2e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 TGACCCCTCTCGACTCGAGAGTTC 27
|||||
Db 368 TGAACCTCTCTGAGTGAAGAGCTCC 343

RESULT 7
BF889056/c 123 bp mRNA linear EST 18-JAN-2001
LOCUS MR4-TN0116-271100-202-e06 TN0116 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF889056
ACCESSION BF889056.1 GI:12280302
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovsky-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.G.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G., Laboratory of Cancer Genetics, Ludwig Institute for Cancer Research, Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR4t2-MR4-TN0116-271100-202-e06&t3=2000-11-27&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 68
High quality sequence stop: 123.
Location/Qualifiers
1..123
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="TN0116"
/note="Organ: testis normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
ORIGIN
Query Match 64.7%; Score 19.4; DB 2; Length 123;
Best Local Similarity 79.3%; Pred. No. 9.6e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 CTGACCCCTCTGAGCTGAGAGTTCCGC 29
41 CTGACGCTCTGAGCTGAGAGTCTCTC 13
Db

SOURCE
ORGANISM Leishmania braziliensis
REFERENCE Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Leishmania; Leishmania braziliensis species complex.
AUTHORS Laurentino, E.C., Ruiz, J.C. and Cruz, A.K.
TITLE GSS analysis of the Leishmania braziliensis genome
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 168)
AUTHORS Cruz, A.K.
TITLE Direct Submission
JOURNAL Submitted (17-JUN-2003) Cruz A.K., University of Sao Paulo, Department of Molecular and Cell Biology, FMRP, Avenida Bandeirantes 3900, Ribeirao Preto, SP, 14049-900, BRAZIL
COMMENT Clone requests: akcruz@fmrp.usp.br.
FEATURES
LOCATION/Qualifiers
1..168
/organism="Leishmania braziliensis"
/mol_type="genomic DNA"
/strain="WHOM/BR/75/M2904"
/db_xref="taxon:5660"
/clone="LBAF041G06"
ORIGIN
Query Match 64.7%; Score 19.4; DB 9; Length 168;
Best Local Similarity 79.3%; Pred. No. 9.6e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 CTGACCCCTCTGAGCTGAGAGTTCCGC 29
43 CTGACCCCTCTGAGCTGAGAGTTCCGC 71
Db

RESULT 9
BI055318/c 240 bp mRNA linear EST 15-JUN-2001
LOCUS RC0-GN0233-060201-031-f07 GN0233 Homo sapiens cDNA, mRNA sequence.
DEFINITION BI055318
ACCESSION BI055318.1 GI:14462848
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovsky-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.G.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G., Laboratory of Cancer Genetics, Ludwig Institute for Cancer Research, Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC0t2-RC0-GN0233-060201-031-f07&t3=2001-02-06&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 69
High quality sequence stop: 156.
Location/Qualifiers
1..240
FEATURES
SOURCE

Db 29 CTGAGCTGCTGACTTAGAGATCCAC 1

RESULT 12
LOCUS B1516790 469 bp mRNA linear EST 08-APR-2002
DEFINITION B160023B10C10.5 Bee Brain Normalized Library, B16 Apis mellifera
ACCESSION CDNA clone B160023B10C10 5', mRNA sequence.
VERSION B1516790
KEYWORDS B1516790.1 GI:15367164
SOURCE EST.
ORGANISM Apis mellifera (honey bee)
Apis mellifera
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea; Apidae; Apis.

REFERENCE
AUTHORS 1 (bases 1 to 469)
TITLE Whitfield, C.W., Band, M.R., Bonaldo, M.F., Kumar, C.G., Liu, L., Pardinas, J., Robertson, H.M., Soares, B. and Robinson, G.E., Annotated expressed sequence tags and cDNA microarrays for studies of brain and behavior in the honey bee
JOURNAL Genome Res. 12 (4), 555-566 (2002)
MEDLINE 21929762
PUBMED 11932240
COMMENT Contact: Gene E. Robinson
Department of Entomology
University of Illinois
505 S. Goodwin Ave., Urbana, IL 61801, USA
Tel: 217 265 0309
Fax: 217 244 3499
Email: generobi@life.uiuc.edu
This research was funded by the University of Illinois Critical Research Initiatives Fund and a Burroughs-Wellcome Trust Innovation Award in Functional Genomics to G.E. Robinson and an NSF postdoctoral fellowship in Bioinformatics to C.W. Whitfield.
PCR Primers
FORWARD: TATATGACTCCTACTATAGG
BACKWARD: ATTACCTCCTACTAAG
Plate: B160023B10 row: C column: 10
Seq primer: AGCGGTACATTTTCACACGGA
High quality sequence stop: 469.
location/Qualifiers
1..469
/organism="Apis mellifera"
/mol_type="mRNA"
/strain="mixed strains of European bees, predominantly A.m. ligustica"
/db_xref="taxon:7460"
/clone="B160023B10C10"
/sex="female"
/tissue_type="brain"
/dev_stage="adult worker honey bee"
/lab_host="DH10B"
/clone_lib="Bee Brain Normalized Library, B16"
/note="Organ: brain; Vector: pTV3-Pac; Site: 1; Ecot1; Soares laboratory and it was constructed and normalized as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from dissected brains of adult worker bees of various ages and various behavioral groups."

FEATURES
source

ORIGIN
Query Match 64.7%; Score 19.4; DB 4; Length 469;
Best Local Similarity 95.2%; Pred. No. 9.9e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 7 CCCTTCGACTCGAGATTC 27
DB 273 CCCTTCGACTCGAGCTTC 293

RESULT 13

B1514171
LOCUS B1514171 511 bp mRNA linear EST 08-APR-2002
DEFINITION B160014B10C05.5 Bee Brain Normalized Library, B16 Apis mellifera
ACCESSION CDNA clone B160014B10C05 5', mRNA sequence.
VERSION B1514171
KEYWORDS B1514171.1 GI:15364545
SOURCE EST.
ORGANISM Apis mellifera (honey bee)
Apis mellifera
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea; Apidae; Apis.

REFERENCE
AUTHORS 1 (bases 1 to 511)
TITLE Whitfield, C.W., Band, M.R., Bonaldo, M.F., Kumar, C.G., Liu, L., Pardinas, J., Robertson, H.M., Soares, B. and Robinson, G.E., Annotated expressed sequence tags and cDNA microarrays for studies of brain and behavior in the honey bee
JOURNAL Genome Res. 12 (4), 555-566 (2002)
MEDLINE 21929762
PUBMED 11932240
COMMENT Contact: Gene E. Robinson
Department of Entomology
University of Illinois
505 S. Goodwin Ave., Urbana, IL 61801, USA
Tel: 217 265 0309
Fax: 217 244 3499
Email: generobi@life.uiuc.edu
This research was funded by the University of Illinois Critical Research Initiatives Fund and a Burroughs-Wellcome Trust Innovation Award in Functional Genomics to G.E. Robinson and an NSF postdoctoral fellowship in Bioinformatics to C.W. Whitfield.
PCR Primers
FORWARD: TATATGACTCCTACTATAGG
BACKWARD: ATTACCTCCTACTAAG
Plate: B160014B10 row: C column: 05
Seq primer: AGCGGTACATTTTCACACGGA
High quality sequence stop: 511.
location/Qualifiers
1..511
/organism="Apis mellifera"
/mol_type="mRNA"
/strain="mixed strains of European bees, predominantly A.m. ligustica"
/db_xref="taxon:7460"
/clone="B160014B10C05"
/sex="female"
/tissue_type="brain"
/dev_stage="adult worker honey bee"
/lab_host="DH10B"
/clone_lib="Bee Brain Normalized Library, B16"
/note="Organ: brain; Vector: pTV3-Pac; Site: 1; Ecot1; Soares laboratory and it was constructed and normalized as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from dissected brains of adult worker bees of various ages and various behavioral groups."

FEATURES
source

ORIGIN
Query Match 64.7%; Score 19.4; DB 4; Length 511;
Best Local Similarity 95.2%; Pred. No. 9.9e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 7 CCCTTCGACTCGAGATTC 27
DB 274 CCCTTCGACTCGAGCTTC 294

RESULT 14
LOCUS B1513884 541 bp mRNA linear EST 08-APR-2002
DEFINITION B160014A10F07.5 Bee Brain Normalized Library, B16 Apis mellifera
CDNA clone B160014A10F07 5', mRNA sequence.

ACCESSION BI513884
 VERSION BI513884.1 GI:15364258
 KEYWORDS
 SOURCE Apis mellifera (honey bee)
 ORGANISM Apis mellifera
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea; Apidae; Apis.
 1 (bases 1 to 541)
 Whitfield,C.W., Band,M.R., Bonaldo,M.F., Kumar,C.G., Liu,L., Pardinas,J., Robertson,H.M., Soares,B. and Robinson,G.E.
 Annotated expressed sequence tags and cDNA microarrays for studies of brain and behavior in the honey bee
 Genome Res. 12 (4), 555-566 (2002)
 21929762
 11932240
 JOURNAL MEDLINE
 PUBMED
 COMMENT
 Contact: Gene E. Robinson
 Department of Entomology
 University of Illinois
 505 S. Goodwin Ave., Urbana, IL 61801, USA
 Tel: 217 265 0309
 Fax: 217 244 3499
 Email: generobi@life.uiuc.edu
 This research was funded by the University of Illinois Critical Research Initiatives Fund and a Burroughs-Wellcome Trust Innovation Award in Functional Genomics to G.E. Robinson and an NSF Postdoctoral Fellowship in Bioinformatics to C.W. Whitfield.
 PCR Primers
 FORWARD: TAATAGCACTCACTATAGCG
 BACKWARD: ATTAACCTCTACTAAG
 Plate: BB160014A11 row: F column: 07
 Seq primer: AGCGGATACAAATTCACACAGA
 High quality sequence stop: 541.
 Location/Qualifiers
 1..541
 /organism="Apis mellifera"
 /mol_type="mRNA"
 /strain="mixed strains of European bees, predominantly A.m. ligustica"
 /db_xref="taxon:7460"
 /clone="BB160014A11F07"
 /sex="female"
 /tissue_type="brain"
 /dev_stage="adult worker honey bee"
 /lab_host="DH10B"
 /clone_1lb="Bee Brain Normalized Library, BB16"
 /note="Organ: Brain; Vector: pRTT3-Pac; Site: 1: Ecot1; Site 2: Not1; The BB16 library was contributed by the Soares laboratory and it was constructed and normalized as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from dissected brains of adult worker bees of various ages and various behavioral groups."

ORIGIN
 Query Match 64.7%; Score 19.4; DB 4; Length 541;
 Best Local Similarity 95.2%; Pred. No. 9.9e+02;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 CCCTCTGAGCTGAGAGTTCC 27
 |||||
 DB 182 CCCTCTGAGCTGAGAGTTCC 202

RESULT 15
 BI513971
 LOCUS BB160014A11F07 5 Bee Brain Normalized Library, BB16 Apis mellifera
 DEFINITION CDNA clone BB160014A11F07 5', mRNA sequence.
 ACCESSION BI513971
 VERSION BI513971.1 GI:15364345
 KEYWORDS EST.
 SOURCE Apis mellifera (honey bee)

ORGANISM Apis mellifera
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea; Apidae; Apis.
 1 (bases 1 to 564)
 Whitfield,C.W., Band,M.R., Bonaldo,M.F., Kumar,C.G., Liu,L., Pardinas,J., Robertson,H.M., Soares,B. and Robinson,G.E.
 Annotated expressed sequence tags and cDNA microarrays for studies of brain and behavior in the honey bee
 Genome Res. 12 (4), 555-566 (2002)
 21929762
 11932240
 JOURNAL MEDLINE
 PUBMED
 COMMENT
 Contact: Gene E. Robinson
 Department of Entomology
 University of Illinois
 505 S. Goodwin Ave., Urbana, IL 61801, USA
 Tel: 217 265 0309
 Fax: 217 244 3499
 Email: generobi@life.uiuc.edu
 This research was funded by the University of Illinois Critical Research Initiatives Fund and a Burroughs-Wellcome Trust Innovation Award in Functional Genomics to G.E. Robinson and an NSF Postdoctoral Fellowship in Bioinformatics to C.W. Whitfield.
 PCR Primers
 FORWARD: TAATAGCACTCACTATAGCG
 BACKWARD: ATTAACCTCTACTAAG
 Plate: BB160014A11 row: F column: 07
 Seq primer: AGCGGATACAAATTCACACAGA
 High quality sequence stop: 564.
 Location/Qualifiers
 1..564
 /organism="Apis mellifera"
 /mol_type="mRNA"
 /strain="mixed strains of European bees, predominantly A.m. ligustica"
 /db_xref="taxon:7460"
 /clone="BB160014A11F07"
 /sex="female"
 /tissue_type="brain"
 /dev_stage="adult worker honey bee"
 /lab_host="DH10B"
 /clone_1lb="Bee Brain Normalized Library, BB16"
 /note="Organ: Brain; Vector: pRTT3-Pac; Site: 1: Ecot1; Site 2: Not1; The BB16 library was contributed by the Soares laboratory and it was constructed and normalized as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from dissected brains of adult worker bees of various ages and various behavioral groups."

ORIGIN
 Query Match 64.7%; Score 19.4; DB 4; Length 564;
 Best Local Similarity 95.2%; Pred. No. 9.9e+02;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 CCCTCTGAGCTGAGAGTTCC 27
 |||||
 DB 182 CCCTCTGAGCTGAGAGTTCC 202

Search completed: March 1, 2005, 01:58:25
 Job time : 2088.63 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2005, 23:26:41; Search time 261.818 Seconds
(without alignments)
678.304 Million cell updates/sec

Title: US-10-086-062-4

Perfect score: 30

Sequence: 1 ctgagaccctctcgcactcagagagtcgcgt 30

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues 8780412

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq.16Dec04:*
1: geneeqn1808:*
2: geneeqn1908:*
3: geneeqn2008:*
4: geneeqn2001as:*
5: geneeqn2001bs:*
6: geneeqn2002as:*
7: geneeqn2002bs:*
8: geneeqn2003as:*
9: geneeqn2003bs:*
10: geneeqn2003cs:*
11: geneeqn2003ds:*
12: geneeqn2004as:*
13: geneeqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	30	100.0	30	AAD24069	Aad24069 Maize eng
2	19.2	64.0	1500	AA511024	Aa511024 Vibrio ch
3	19	63.3	121124	AD97107	Ad97107 Mouse can
4	18.8	62.7	2461	AB741766	Ab741766 Toxicity
5	18.8	62.7	23107	9 ADA02762	Ada02762 Human RUN
6	18.8	62.7	23107	10 ADC85242	Adc85242 Human Run
7	18.8	62.7	23107	12 ADM74357	Adm74357 Human car
8	18.8	62.7	23107	12 ADM74357	Adm74357 Human car
9	18.6	62.0	3210	AA227624	Aa227624 Plasmid S
10	18.6	62.0	3459	AA227623	Aa227623 Plasmid S
11	18.4	61.3	319608	3 AAH51601	Aah51601 Human chr
12	18.4	61.3	319608	5 AAS09301	Aas09301 Human chr
13	18.2	60.7	340449	8 AAL52198	Aal52198 Human sec
14	18	60.0	42325	10 ADB74382	Adb74382 Mycobacte
15	18	60.0	110000	11 ADM27081_10	Adm27081_10 O
16	17.8	59.3	165	10 AAD49400	Ad49400 Human
17	17.8	59.3	254	12 ADG00322	Adg00322 Nicotiana
18	17.8	59.3	369	11 ABD11207	Abd11207 Pseudomon
19	17.8	59.3	426	10 ADD49385	Add49385 Human lun
20	17.8	59.3	449	10 ADD49343	Add49343 Human lun

21	17.8	59.3	449	10	ADD49294	Add49294 Human lun
22	17.8	59.3	986	6	ABQ46720	Abq46720 Oligonuc1
23	17.8	59.3	986	6	ABQ46721	Abq46721 Oligonuc1
24	17.8	59.3	1553	12	ADL12873	Adl12873 Human etc
25	17.8	59.3	1603	4	AAK77077	Aak77077 Human lmm
26	17.8	59.3	1731	11	ACN91401	Acn91401 Breast ca
27	17.8	59.3	2780	13	ADR25834	Adr25834 Breast ca
28	17.8	59.3	152141	8	ACA64961	Ac64961 Human BCR
29	17.8	59.3	15395	13	ADP33534	Adp33534 Murine ca
30	17.6	58.7	246	12	ADQ52469	Adq52469 Human met
31	17.6	58.7	8119	3	AA235392	Aa235392 Maize etc
32	17.6	58.7	49243	4	ABLO3188	Ablo3188 Drosoph11
33	17.4	58.0	65	6	ABN51822	Abn51822 Mouse spl
34	17.4	58.0	300	2	AA213036	Aa213036 Human gen
35	17.4	58.0	300	2	AA298464	Aa298464 Human can
36	17.4	58.0	411	4	AA185652	Aa185652 Human pol
37	17.4	58.0	460	9	ACH41661	Ach41661 Human toe
38	17.4	58.0	594	12	ACH74968	Ach74968 Human toe
39	17.4	58.0	699	2	AA215926	Aa215926 Human gen
40	17.4	58.0	729	12	ADO63441	Ado63441 Transcrip
41	17.4	58.0	778	6	ABK30430	Abk30430 Human G-P
42	17.4	58.0	936	13	ADR92877	Adr92877 Novel S.
43	17.4	58.0	949	13	ADS50577	Ad50577 Bacterial
44	17.4	58.0	1210	4	AA541038	Aa541038 CDNA enco
45	17.4	58.0	1210	4	AA534819	Aa534819 CDNA enco

ALIGNMENTS

RESULT 1	
AAD24069	
ID	AAD24069 standard; DNA; 30 BP.
XX	
AC	AAD24069;
XX	
DT	09-APR-2002 (first entry)
XX	
DE	Maize engineered Ubi-1 promoter heat shock element #3.
XX	
KW	Gene expression; maize; ubiquitin promoter; Ubi-1; HSE;
KW	heat shock element; agronomic gene; dg.
XX	
OS	Zea mays.
XX	
EH	Key
FT	misc_feature
FT	Location/Qualifiers
FT	1..15
FT	/tag= "a"
FT	/note= "5' heat shock element"
FT	16..30
FT	/tag= "b"
FT	/note= "3' heat shock element"
XX	
PN	WO200194394-A2.
XX	
PD	13-DEC-2001.
XX	
PF	08-JUN-2001; 2001WO-US018689.
XX	
PR	09-JUN-2000; 2000US-00530558.
XX	
PA	(PROD-) PRODIGENE INC.
XX	
PI	Jilka JM, Hood EE, Howard JA;
XX	
DR	WPI; 2002-122117/16.
XX	
PT	New promoter sequences for causing expression of a structural gene
PT	especially agronomic gene or open reading frame in a plant cell,
XX	comprises engineered versions of the maize ubiquitin promoter.
PS	Claim 6; Page 54; 68bp; English.
XX	

CC The invention relates to a promoter sequence capable of directing
CC expression of a nucleotide sequence in a plant cell, comprising maize
CC ubiquitin (ubi-1) promoter sequence with a modification so that it does
CC not include two overlapping heat shock elements (HSE) or it directs
CC expression to increase the endosperm/embryo expression ratio of the
CC protein when compared to the ratio from a wild-type ubiquitin promoter.
CC The modified ubi-1 promoter comprises a deletion of 3', 5' or both HSEs,
CC two non-overlapping/adjacent HSEs, replacement of HSEs with a trimer of a
CC seed specific element from the promoter of pea lectin gene *Pel*, or
CC insertion of a transcription factor binding site in the HSE region. An
CC expression construct comprising modified ubi-1 promoter is useful for
CC causing expression of a structural gene (agronomic genes) or open reading
CC frame in a plant cell. The modified ubi-1 promoter increases expression
CC levels beyond those observed with native ubiquitin promoter. The present
CC sequence is maize engineered ubi-1 promoter with heat shock elements
CC adjacently placed. Note: The present sequence is also shown in claim 26,
CC page 56 of the specification. However, this sequence has an additional
CC nucleotide at the 3' end
XX

SO Sequence 30 BP; 4 A; 12 C; 7 G; 7 T; 0 U; 0 Other;

Query Match 100.0%; Score 30; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACCCCTCTCGACTCGAGATTCCGCT 30
DB 1 CTGGACCCCTCTCGACTCGAGATTCCGCT 30

RESULT 2

AAS11024/c
ID AAS11024 standard; DNA; 1500 BP.

AC AAS11024;
DT 11-SEP-2003 (revised)
DT 24-OCT-2001 (first entry)

XX Vibrio cholera 16S ribosomal RNA gene.

KM Antisense; bacterial 16S ribosomal RNA; rRNA; bacterial infection; human;
KM food grain supplement; livestock; poultry; therapeutic; ds.

OS Vibrio cholerae.

PN WO200142457-A2.

PD 14-JUN-2001.

PF 29-NOV-2000; 2000WO-US042391.

PR 29-NOV-1999; 99US-0168150P.

PA (AVIB-) AVI BIOPHARMA INC.

PI Iversen PL;

DR WPI; 2001-457295/49.

PT Antibacterial compound, useful for treating bacterial infections and as
PT livestock and poultry food supplement, comprises antisense
PT oligonucleotides complementary to bacterial 16S and 23S rRNA.

PS Disclosure; Page; 62pp; English.

CC AAS11021-AAS11034 represent the coding sequences of bacterial 16S
CC ribosomal RNA (rRNA) genes. The sequences were used to design anti-
CC bacterial compounds comprising substantially uncharged antisense
CC oligomers containing 8-40 nucleotide subunits, including a targeting
CC nucleic acid sequence at least 10 nucleotides in length which is
CC complementary to a bacterial 16S or 23S rRNA nucleic acid sequence. The
CC antisense oligomers are used for treating a bacterial infection in a

CC human or a mammalian animal produced by *Escherichia coli*, *Salmonella*
CC typhimurium, *Pseudomonas aeruginosa*, *Vibrio cholera*, *Neisseria*
CC gonorrhoea, *Helicobacter pylori*, *Bartonella henselae*, *Haemophilus*
CC influenza, *Shigella dysenteriae*, *Staphylococcus aureus*, *Mycobacterium*
CC tuberculosis, *Streptococcus pneumoniae*, *Treponema pallidum* and *Chlamydia*
CC trachomatis. The antibacterial compound may be used as a food grain
CC supplement in livestock and poultry food composition. Note: The present
CC sequence is not shown in the specification but has been accessed from
CC Genbank using the appropriate accession number given in the
CC specification. (Updated on 11-SEP-2003 to standardise OS field)
XX

SO Sequence 1500 BP; 376 A; 326 C; 482 G; 312 T; 0 U; 4 Other;

Query Match 64.0%; Score 19.2; DB 5; Length 1500;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 21; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CTGGACCCCTCTCGACTCGAGATTCCG 28
DB 212 CTGGACCCATCCGACGAFARATCCG 185

RESULT 3

ADQ97107/c
ID ADQ97107 standard; DNA; 121124 BP.

AC ADQ97107;

DT 07-OCT-2004 (first entry)

DE Mouse cancer associated sequence MD08-002, SEQ ID 83.

KM Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.

OS Mus musculus.

PN WO2004060304-A2.

PD 22-JUL-2004.

PF 22-DEC-2003; 2003WO-US041389.

PR 27-DEC-2002; 2002US-00330773.

PA (SAGR-) SAGRES DISCOVERY INC.

PI Morris DW, Malandro MS;

DR WPI; 2004-543781/52.

PT New isolated cancer associated nucleic acids comprising at least 10
PT contiguous nucleotides, useful for diagnosing, preventing and/or treating
PT cancers such as leukemia and lymphoma.

PS Claim 1; SEQ ID NO 83; 199pp; English.

CC The present invention relates to cancer associated sequences (ADQ97025-
CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 121124 BP; 32972 A; 25314 C; 26641 G; 35451 T; 0 U; 746 Other;

Query Match 63.3%; Score 19; DB 12; Length 121124;
Best Local Similarity 81.5%; Pred. No. 1.9e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 GACCCCTCTCGACTCGAGATTCCGCT 30
DB 88764 GCCCCTCTCGATTACAGATTACGCT 88738

RESULT 4
 ABT41766/C
 ID ABT41766 standard; DNA; 2481 BP.
 XX
 AC ABT41766;
 XX
 DT 26-JUN-2003 (first entry)
 XX
 DE Toxicity modelling related rat gene SEQ ID NO 1468.
 XX
 KW Toxic effect; gene expression profile; renal toxicity; toxicity marker;
 KW database; drug screening; toxicity assay; rat; ds.
 XX
 OS Rattus norvegicus.
 XX
 PN WO200295000-A2.
 XX
 PD 28-NOV-2002.
 XX
 PF 22-MAY-2002; 2002WO-US016173.
 XX
 PR 22-MAY-2001; 2001US-0292335P.
 PR 13-JUN-2001; 2001US-0297523P.
 PR 19-JUN-2001; 2001US-0298925P.
 PR 10-JUL-2001; 2001US-0303807P.
 PR 10-JUL-2001; 2001US-0303808P.
 PR 10-JUL-2001; 2001US-0303810P.
 PR 28-AUG-2001; 2001US-0315047P.
 PR 27-SEP-2001; 2001US-0324928P.
 PR 22-OCT-2001; 2001US-0330462P.
 PR 01-NOV-2001; 2001US-0330867P.
 PR 21-NOV-2001; 2001US-0331805P.
 PR 06-DEC-2001; 2001US-0336144P.
 PR 19-DEC-2001; 2001US-0340873P.
 PR 21-FEB-2002; 2002US-0357842P.
 PR 21-FEB-2002; 2002US-0357843P.
 PR 21-FEB-2002; 2002US-0357844P.
 PR 15-MAR-2002; 2002US-0364134P.
 PR 08-APR-2002; 2002US-0370144P.
 PR 08-APR-2002; 2002US-0370206P.
 PR 08-APR-2002; 2002US-0370247P.
 PR 17-APR-2002; 2002US-0372794P.
 PR 21-APR-2002; 2002US-0371679P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
 XX
 DR WPI; 2003-148464/14.
 XX
 DT Predicting at least one toxic effect of a compound, useful for toxicity
 PT modelling, comprises preparing a gene expression profile of a tissue or
 PT cell sample exposed to the compound, and comparing the gene expression
 PT profile to a database.
 XX
 XX Example 4; Page; 446pp; English.
 PS
 XX The invention relates to a novel method of predicting at least one toxic
 XX effect of a compound. The method comprises a gene expression profile of a
 CC tissue or cell sample exposed to the compound, and comparing the gene
 CC expression profile to a database comprising at least part of the data or
 CC information given in the specification. The methods are useful for
 CC predicting at least one toxic effect of a compound, predicting the renal
 CC progression of a toxic effect of a compound, predicting the renal
 CC toxicity of a compound, or identifying toxicity markers in tissues or
 CC cells exposed to known renal toxin. The genes are useful as toxicity
 CC markers in drug screening and toxicity assays, in monitoring disease or
 CC physiological states, or disease progression. This polynucleotide
 CC represents a rat DNA sequence relating to the toxic effect database
 CC described in the specification. NOTE: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from the World Intellectual Property

CC Organization
 XX
 XX Sequence 2481 BP; 539 A; 653 C; 555 G; 674 T; 0 U; 0 Other;
 DB Query Match 62.7%; Score 18.8; DB 10; Length 2481;
 DB Best Local Similarity 76.7%; Pred. No. 1.7e+02;
 DB Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 DY 1 CTGCACCCCTCTCGAGTCCGCT 30
 DB 1077 CTGCACCCCTCTCGTACGAGTTCAGCT 1048
 RESULT 5
 ADA02762
 ID ADA02762 standard; DNA; 23107 BP.
 XX
 AC ADA02762;
 XX
 DT 06-NOV-2003 (first entry)
 XX
 DE Human RUNX3 carcinoma associated gene, SEQ ID NO:1280.
 XX
 KW Human; carcinoma associated; oncogene; carcinoma; cancer; breast;
 KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO2003057146-A2.
 XX
 PD 17-JUL-2003.
 XX
 PF 26-DEC-2002; 2002WO-US041414.
 XX
 PR 26-DEC-2001; 2001US-00035832.
 XX
 PA (SAGR-) SAGRES DISCOVERY.
 XX
 PI Morris DW;
 XX
 DR WPI; 2003-587068/55.
 XX
 DT New recombinant nucleic acid encoding carcinoma associated protein,
 PT useful for preparing compositions for treating carcinomas.
 PT
 XX
 PS Claim 1; SEQ ID NO 1280; 245pp; English.
 XX
 XX The invention relates to recombinant carcinoma associated (CA) nucleic
 CC acid sequences from mouse and human (ADA01482-ADA03094), and to
 CC recombinant carcinoma associated proteins (CAP) encoded by them. The
 CC invention also encompasses expression vectors and host cells comprising a
 CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
 CC binds to the protein, and a blotchip comprising CA nucleic acid or
 CC fragments thereof. The sequences of the invention were identified using
 CC oncogenic retroviruses, which insert into the genome of the host organism
 CC at random. Many of these do not carry transduced host oncogenes or
 CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
 CC direct consequence of the effects of proviral integration into host
 CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
 CC carcinoma (especially breast cancer, prostate cancer, lymphoma or
 CC leukaemia) or a propensity to carcinoma by determination of the sequence
 CC of a CA gene, or by determination of CA gene expression in particular
 CC tissues. CA nucleic acids, proteins and antibodies are also useful as
 CC therapeutic agents and in screening and evaluating drug candidates. The
 CC present sequence represents a specifically claimed human CA nucleic acid
 CC sequence of the invention. Note: The complete sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 23107 BP; 5308 A; 6264 C; 6229 G; 5306 T; 0 U; 0 Other;

Query Match 62.7%; Score 18.8; DB 9; Length 23107;
Best Local Similarity 76.7%; Pred. No. 2.1e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CTGAGACCCCTCTGACTCGAGAGTTCCGCT 30
DB 14873 CTGGAGACCCCTCTGACTCGAGAGTTCCGCT 14902

RESULT 6
ADB72500
ID ADB72500 standard; DNA; 23107 BP.

XX ADB72500;

DT 04-DEC-2003 (first entry)

DE Human Runx3 gene.

KW human; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.

OS Homo sapiens.

PN WO2003008583-A2.

PD 30-JAN-2003.

PF 26-DEC-2001; 2001WO-US051291.

PR 02-MAR-2001; 2001US-00798586.

PR 23-OCT-2001; 2001US-00004113.

PR 08-NOV-2001; 2001US-00052482.

PR 30-NOV-2001; 2001US-00997722.

PR 20-DEC-2001; 2001US-00034650.

PA (SAGR-) SAGRES DISCOVERY.

PI Morris DW, Engelhard EK;

DR WPI; 2003-239337/23.

PT New recombinant nucleic acid, useful for treating carcinomas, lymphomas,

XX cancer; neoplasm, adenocarcinoma, or sarcomas.

PS Claim 1; SEQ ID NO 328; 2304bp; English.

CC The invention relates to a novel recombinant nucleic acid comprising a

CC nucleotide sequence selected from any of the 660 sequences fully defined

CC in the specification. A polynucleotide of the invention has cytostatic

CC activity, and may have a use in gene therapy, or in a vaccine. The

CC recombinant nucleic acids and polypeptides are useful for treating

CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and

CC sarcomas. The present sequence represents a human gene of the invention.

XX Sequence 23107 BP; 5308 A; 6264 C; 6229 G; 5306 T; 0 U; 0 Other;

QY Query Match 62.7%; Score 18.8; DB 10; Length 23107;

Best Local Similarity 76.7%; Pred. No. 2.1e+02;

Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CTGAGACCCCTCTGACTCGAGAGTTCCGCT 30

DB 14873 CTGGAGACCCCTCTGACTCGAGAGTTCCGCT 14902

RESULT 7

ADCB5242

ID ADCB5242 standard; DNA; 23107 BP.

XX ADCB5242;

AC ADCB5242;

XX 01-JAN-2004 (first entry)

DT

XX Human Runx3 genomic sequence.

DE Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA;

XX secreted; transmembrane; intracellular; ds.

OS Homo sapiens.

PN WO2003045230-A2.

PD 05-JUN-2003.

PF 02-DEC-2002; 2002WO-US038582.

PR 30-NOV-2001; 2001US-00997722.

PA (SAGR-) SAGRES DISCOVERY.

PI Morris DW, Engelhard EK;

DR WPI; 2003-513603/48.

PT New recombinant nucleic acid comprising a nucleotide sequence of any of

PS the carcinoma-associated (CA) genes, useful for screening for drug

XX candidates for diagnosing or treating carcinomas.

XX Claim 1; SEQ ID NO 28; 983bp; English.

CC The invention relates to a recombinant nucleic acid comprising a

CC nucleotide sequence selected from any of the fully defined carcinoma-

CC associated (CA) genes from the 50 tables given in the specification. The

CC CA proteins are secreted, transmembrane or intracellular proteins. The

CC recombinant nucleic acids are useful for screening for drug candidates

CC for diagnosing or treating carcinomas. Sequences given in ADCB5215-

CC ADCB5514 represent CA genes of the invention.

XX Sequence 23107 BP; 5308 A; 6264 C; 6229 G; 5306 T; 0 U; 0 Other;

QY Query Match 62.7%; Score 18.8; DB 10; Length 23107;

Best Local Similarity 76.7%; Pred. No. 2.1e+02;

Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CTGAGACCCCTCTGACTCGAGAGTTCCGCT 30

DB 14873 CTGGAGACCCCTCTGACTCGAGAGTTCCGCT 14902

RESULT 8

ADM74357

ID ADM74357 standard; DNA; 23107 BP.

XX ADM74357;

AC ADM74357;

XX 01-JUN-2004 (first entry)

DT Human carcinoma associated (CA) nucleic acid #13.

XX Human; carcinoma associated nucleic acid; CA nucleic acid; gene; ds;

XX carcinoma associated protein; CAP; carcinoma; leukemia; lymphoma;

XX cytostatic.

OS Homo sapiens.

PN US2004072154-A1.

PD 15-APR-2004.

PF 30-NOV-2001; 2001US-00997722.

PR 22-DEC-2000; 2000US-00747377.

PR 02-MAR-2001; 2001US-00798586.

XX (MORR/) MORRIS D W.

XX Robert LS, Gledie S;
 XX
 PI MPI, 1999-591104/50.
 XX
 PT Protein expression in floral cells for peptide display, mediating plant
 XX sterility, and modifying pollen-pistil interactions.
 XX
 PS Example 5; Page 100-102; 113pp; English.

CC This sequence represents the plasmid SPF-1, containing a fusion of the
 CC Brassica napus GPR363 gene to the Staphylococcus aureus protease gene. The
 CC invention relates to a method for modifying the extracellular compartment
 CC of a floral cell of a plant, that comprises expressing a construct
 CC comprising a gene of interest encoding a protein, fusion protein or
 CC peptide, or a fragment of them, which is capable of modifying the
 CC composition of the extracellular compartment of the floral cell and
 CC altering either the function, use or development of the floral cell or
 CC modifying the interaction of the floral cell with other cells, within an
 CC anther or pistil cell. The method is used to modify pollen-pistil
 CC interaction or function, which mediates, produces or prevents self-
 CC compatibility, self-incompatibility, out- or in-crossing or combinations
 CC of these. The method is also used for localizing proteins on the surface
 CC of pollen for the purpose of peptide display. The protein localized on
 CC the surface of the pollen may be an antibody or antigen or is a protein
 CC that is effective in controlling insect growth, behaviour, feeding,
 CC development or reproduction

XX Sequence 3459 BP; 1082 A; 721 C; 678 G; 978 T; 0 U; 0 Other;

Query Match 62.0%; Score 18.6; DB 2; Length 3459;
 Best Local Similarity 84.0%; Pred. No. 2.2e+02;

Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 CTGACCCCTCTCGACTCGAGATT 25
 Db 1232 CTGACCCCTCTCGAGATTGATT 1208

RESULT 11

AAH51601 standard; DNA; 319608 BP.

AAH51601;

29-AUG-2001 (first entry)

Human chromosome 13q31-q33 genomic nucleotide sequence.

XX sbg1; g34665; sbg2; g35017; g35018; chromosome 13q31-q33; haplotype;
 KW biallelic marker; polymorphism; schizophrenia; bipolar disorder; ds.
 XX
 OS Homo sapiens.

PN W0200058510-A2.

XX 05-OCT-2000.

PF 30-MAR-2000; 2000MO-IB000435.

XX 30-MAR-1999; 99US-0126903P.
 PR 30-APR-1999; 99US-0131971P.
 PR 14-APR-1999; 99US-0132065P.
 PR 27-JUL-1999; 99US-0143928P.
 PR 29-JUL-1999; 99US-0145915P.
 PR 29-JUL-1999; 99US-0146452P.
 PR 28-OCT-1999; 99US-0146453P.
 PR 28-OCT-1999; 99US-0162288P.

XX (GENSET) GENSET.

PI Cohen D, Blumenfeld M, Chumakov I, Bouguenel L, Bihain B,
 PI Esseloux L;

XX MPI; 2000-619082/59.

XX Polynucleotides comprising sequences from sbg1 and g35018 biallelic
 PT markers are used for genotyping and detecting schizophrenia or bipolar
 PT disorder and predisposition to these disorders.
 XX
 PS Claim 1; Page 409-493; 737pp; English.

CC AAH51601 represents a human genomic nucleotide sequence comprising sbg1,
 CC g34665, sbg2, g35017 and g35018 nucleotide acid sequences located on the
 CC human chromosome 13q31-q33 locus. The nucleotide sequences contain
 CC biallelic markers and polymorphisms. Sequences AAH51602 - AAH51626 and
 CC AAH62907 - AAH62915 represent cDNA human sbg1 cDNA sequences and protein
 CC products. AAH51627 - AAH51631 and AAH62916 - AAH62918 represent g35018
 CC cDNA sequences and protein products. Primers AAH51632 - AAH51699 are used
 CC to isolate sbg1 cDNAs, while sbg1 exons from different primates are
 CC represented by sequences AAH51642 - AAH51699. Nucleotide sequences of
 CC amplicons which comprise biallelic markers located on the chromosome
 CC 13q31-q33 locus are represented in AAH51700 - AAH51817. Biallelic markers
 CC are represented in the sequences by degenerate/undefined base codes. PCR
 CC primers AAH51818 and AAH51819 are used in the isolation of sequences of
 CC the invention. The biallelic marker containing nucleotide sequences are
 CC used to determine the identity of the nucleotide at a biallelic marker in
 CC a sample DNA sequence. The nucleotide sequences may be labelled and used
 CC for genotyping by determining the identity of a nucleotide at a Region D-
 CC related biallelic marker in a biological sample from single or multiple
 CC subjects. By determining the frequency of a nucleotide at a Region D-
 CC population an association between a genotype and a trait, a haplotype and
 CC a trait and a phenotype and a trait can be detected. The sequences can be
 CC used to determine a predisposition to or early onset of schizophrenia or
 CC bipolar disorder or a beneficial response to or side effects related to
 CC treatment against schizophrenia or bipolar disorder

XX Sequence 319608 BP; 101600 A; 56677 C; 58335 G; 102722 T; 0 U; 274 Other;

Query Match 61.3%; Score 18.4; DB 3; Length 319608;
 Best Local Similarity 78.6%; Pred. No. 3.8e+02;

Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CTGACCCCTCTCGACTCGAGATTCCG 28
 Db 8690 CTGAACCATCTCGATTGAGATTACG 8717

RESULT 12

AAH509301 standard; DNA; 319608 BP.

AAH509301;

26-SEP-2001 (first entry)

Human schizophrenia associated gene g35030 and biallelic markers A1-A71.

XX Human; g35030; biallelic marker; A1-A71; chromosome 13q31-q33;
 KW schizophrenia; bipolar disorder; ds.
 XX
 OS Homo sapiens.

XX Key

PH Location/Qualifiers
 FT primer_bind 7938..7958
 FT /*tag= a

FT primer_bind 8297..8315
 FT /*tag= b
 FT /*tag= b
 FT /*tag= b
 FT /*tag= b

FT primer_bind 8304..8328
 FT /*tag= c
 FT /*tag= c
 FT /*tag= c
 FT /*tag= c

FT primer_bind 8316
 FT /*tag= d
 FT /*tag= d
 FT /*tag= d
 FT /*tag= d

FT primer_bind 8316
 FT /*tag= d
 FT /*tag= d
 FT /*tag= d
 FT /*tag= d

```
FT primer_bind complement(8317. .8335)
FT /tag= e
FT /note= "Binds primer 99-27943-150.mis complement"
FT primer_bind complement(8446. .8463)
FT /tag= f
FT /note= "Binds primer 99-27943.pu complement"
FT primer_bind complement(21365. .21385)
FT /tag= g
FT /note= "Binds primer 99-27935.rp"
FT primer_bind complement(21653. .21671)
FT /tag= h
FT /note= "Binds primer 99-27935-193.mis"
FT misc_binding complement(21660. .21684)
FT /tag= i
FT /bound_moiety= "Probe_99-27935-193"
FT misc_feature complement(21672)
FT /tag= j
FT /note= "Biallelic marker A2"
FT primer_bind complement(21673. .21691)
FT /tag= k
FT /note= "Binds primer 99-27935-193.mis complement"
FT primer_bind complement(21845. .21864)
FT /tag= l
FT /note= "Binds primer 99-27935.pu complement"
FT primer_bind complement(65463. .65471)
FT /tag= m
FT /note= "Binds primer 8-128.pu"
FT primer_bind complement(65466. .65484)
FT /tag= n
FT /note= "Binds primer 8-128-33.mis"
FT misc_binding complement(65473. .65497)
FT /tag= o
FT /bound_moiety= "Probe_8-128-33"
FT misc_feature complement(65485)
FT /tag= p
FT /note= "Biallelic marker A3"
FT primer_bind complement(65486. .65504)
FT /tag= q
FT /note= "Binds primer 8-128-33.mis complement"
FT primer_bind complement(65856. .65874)
FT /tag= r
FT /note= "Binds primer 8-128.rp complement"
FT primer_bind complement(95034. .95053)
FT /tag= s
FT /note= "Binds primer 99-31960.pu"
FT primer_bind complement(95377. .95395)
FT /tag= t
FT /note= "Binds primer 99-31960-363.mis"
FT misc_binding complement(95384. .95408)
FT /tag= u
FT /bound_moiety= "Probe_99-31960-363"
FT misc_feature complement(95396)
FT /tag= v
FT /note= "Biallelic marker A4"
FT primer_bind complement(95397. .95415)
FT /tag= w
FT /note= "Binds primer 99-31960-363.mis complement"
FT primer_bind complement(95543. .95563)
FT /tag= x
FT /note= "Binds primer 99-31960.rp complement"
FT primer_bind complement(107022. .107040)
FT /tag= y
FT /note= "Binds primer 99-24656.pu"
FT primer_bind complement(107262. .107280)
FT /tag= z
FT /note= "Binds primer 99-24656-260.mis"
FT misc_binding complement(107269. .107293)
FT /tag= aa
FT /bound_moiety= "Probe_99-24656-260"
FT misc_feature complement(107281)
FT /tag= ab
FT /note= "Biallelic marker A5"
FT primer_bind complement(107282. .107300)

FT primer_bind /tag= ac
FT /note= "Binds primer 99-24656-260.mis complement"
FT primer_bind complement(107495. .107513)
FT /tag= ad
FT /note= "Binds primer 99-24656.rp complement"
FT primer_bind complement(160279. .160298)
FT /tag= ae
FT /note= "Binds primer 99-24639.rp"
FT primer_bind complement(160621. .160639)
FT /tag= af
FT /note= "Binds primer 99-24639-163.mis"
FT misc_binding complement(160628. .160652)
FT /tag= ag
FT /bound_moiety= "Probe_99-24639-163"
FT misc_feature complement(160640)
FT /tag= ah
FT /note= "Biallelic marker A6"
FT primer_bind complement(160641. .160659)
FT /tag= ai
FT /note= "Binds primer 99-24639-163.mis complement"
FT primer_bind complement(160770. .160787)
FT /tag= aj
FT /note= "Binds primer 99-24634.pu"
FT primer_bind complement(160785. .160802)
FT /tag= ak
FT /note= "Binds primer 99-24639.pu complement"
FT primer_bind complement(160857. .160875)
FT /tag= al
FT /note= "Binds primer 99-24634-108.mis"
FT misc_binding complement(160864. .160888)
FT /tag= am
FT /bound_moiety= "Probe_99-24634-108"
FT misc_feature complement(160876)
FT /tag= an
FT /note= "Biallelic marker A7"
FT primer_bind complement(160877. .160895)
FT /tag= ao
FT /note= "Binds primer 99-24634-108.mis complement"
FT primer_bind complement(161240. .161257)
FT /tag= ap
FT /note= "Binds primer 99-24634.rp complement"
FT primer_bind complement(168813. .168830)
FT /tag= aq
FT /note= "Binds primer 99-7652.pu"
FT primer_bind complement(168955. .168973)
FT /tag= ar
FT /note= "Binds primer 99-7652-162.mis"
FT misc_binding complement(168962. .168986)
FT /tag= as
FT /bound_moiety= "Probe_99-7652-162"
FT misc_feature complement(168974)
FT /tag= at
FT /note= "Biallelic marker A8"
FT primer_bind complement(168975. .168993)
FT /tag= au
FT /note= "Binds primer 99-7652-162.mis complement"
FT primer_bind complement(169331. .169351)
FT /tag= av
FT /note= "Binds primer 99-7652.rp complement"
FT primer_bind complement(170666. .170686)
FT /tag= aw
FT /note= "Binds primer 99-16100.pu"
FT primer_bind complement(170791. .170809)
FT /tag= ax
FT /note= "Binds primer 99-16100-147.mis"
FT misc_binding complement(170798. .170822)
FT /tag= ay
FT /bound_moiety= "Probe_99-16100-147"
FT misc_feature complement(170810)
FT /tag= az
FT /note= "Biallelic marker A9"
FT primer_bind complement(170811. .170829)
FT /tag= ba
```

```
FT primer_bind /note= "Binds primer 99-16100-147.mis complement"
FT complement(171153..171173)
FT /tag= bb
FT /note= "Binds primer 99-16100.rp complement"
FT primer_bind 173065..173085
FT /tag= bc
FT /note= "Binds primer 99-5862.rp"
FT primer_bind 173339..173357
FT /tag= bd
FT /note= "Binds primer 99-5862-167.mis"
FT misc_binding 173346..173370
FT /tag= be
FT /bound_molecy= "Probe_99-5862-167"
FT misc_feature 173358
FT /tag= bf
FT /note= "Biallelic marker A10"
FT primer_bind complement(173359..173377)
FT /tag= bg
FT /note= "Binds primer 99-5862-167.mis complement"
FT primer_bind complement(173495..173514)
FT /tag= bh
FT /note= "Binds primer 99-5862.pu complement"
FT primer_bind 189753..189771
FT /tag= bi
FT /note= "Binds primer 99-5919.pu"
FT primer_bind 189938..189956
FT /tag= bj
FT /note= "Binds primer 99-5919-215.mis"

Query Match 61.3%; Score 18.4; DB 5; Length 319608;
Best Local Similarity 78.6%; Pred No. 3.8e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTGAGCCCTCTCGACTCGAGATTCCG 28
Db 8690 CTGAGCCATCTCGATTGAGATTACG 8717

RESULT 13
AAL52198
ID AAL52198 standard; cDNA; 340449 BP.
XX
XX AAL52198;
XX
XX 22-SEP-2003 (first entry)
XX
XX Human secreted protein genomic DNA coding sequence.
XX
XX Human; gene; ds; secreted protein; chromosome 5; tissue typing;
XX secreted protein-related disease; transgenic animal; drug screening;
XX pharmacogenomic analysis; single nucleotide polymorphism; SNP.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX variation replace(1107, T)
FT /tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(5479, G)
FT /tag= b
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(5480, G)
FT /tag= c
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(6100, T)
FT /tag= d
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(6122, A)
FT /tag= e
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(6683, G)
FT /tag= f
FT /standard_name= "Single nucleotide polymorphism"

FT variation replace(7641, G)
FT /tag= g
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(8493, T)
FT /tag= h
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(9626, A)
FT /tag= i
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(9796, A)
FT /tag= j
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(9868, A)
FT /tag= k
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(10209, C)
FT /tag= l
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(10311, G)
FT /tag= m
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(10831, T)
FT /tag= n
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(12230, C)
FT /tag= o
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(13378, C)
FT /tag= p
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(13423, T)
FT /tag= q
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(13425, G)
FT /tag= r
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(14116, G)
FT /tag= s
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(14508, C)
FT /tag= t
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(14645, A)
FT /tag= u
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(14932, A)
FT /tag= v
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(16272, G)
FT /tag= w
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(16945, T)
FT /tag= x
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(17065, G)
FT /tag= y
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(17442, G)
FT /tag= z
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(17475, C)
FT /tag= aa
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(17844, A)
FT /tag= ab
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(17929, G)
FT /tag= ac
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(18540, C)
FT /tag= ad
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(18592, T)
```



```

FT      /tag= ae
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(18616, G)
FT      /tag= af
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(20491, T)
FT      /tag= ag
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(22023, G)
FT      /tag= ah
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(22248, G)
FT      /tag= ai
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(23118, C)
FT      /tag= aj
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(23789, A)
FT      /tag= ak
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(23800, A)
FT      /tag= al
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(23930, A)
FT      /tag= am
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(24159, G)
FT      /tag= an
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(26339, A)
FT      /tag= ao
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(26508, G)
FT      /tag= ap
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(26570, A)
FT      /tag= aq
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(26582, G)
FT      /tag= ar
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(26693, A)
FT      /tag= as
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(26884, T)
FT      /tag= at
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(27320, T)
FT      /tag= au
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(27339, A)
FT      /tag= av
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(27542, G)
FT      /tag= aw
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(28586, A)
FT      /tag= ax
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(28591, T)
FT      /tag= ay
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(28599, A)
FT      /tag= az
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(30857, A)
FT      /tag= ba
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(31503, T)
FT      /tag= bb
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(33671, A)
FT      /tag= bc

```

```

FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(35045, C)
FT      /tag= bd
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(35944, G)
FT      /tag= be
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(37157, G)
FT      /tag= bf
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(37603, T)
FT      /tag= bg
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(39242, G)
FT      /tag= bh
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(39404, A)
FT      /tag= bi
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(40395, C)
FT      /tag= bj

```

```

Query Match      60.7%; Score 18.2; DB 8; Length 340449;
Best Local Similarity 87.0%; Pred. No. 4.7e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY      2 TGGACCCCTTCGACTCGAGACT 24
DB      25999 TGGACCCCTCCCACTCGAGACT 26021

```

```

RESULT 14
ADB74382
ID      ADB74382 standard; DNA; 42325 BP.
AC      ADB74382;
XX      XX
DT      04-DEC-2003 (first entry)
XX      XX
DE      Mycobacterium leprae DNA #16.
XX      XX
KW      Non-naturally occurring peptide; anion pump protein; tuberculosis;
KW      hypersensitivity reaction; tuberculostatic; gene; ds.
XX      XX
OS      Mycobacterium leprae.
XX      XX
PN      US6583266-B1.
XX      XX
PD      24-JUN-2003.
XX      XX
PF      16-SEP-1994; 94US-00311731.
XX      XX
PR      19-AUG-1993; 93US-00109181.
PR      22-OCT-1993; 93US-00142558.
XX      XX
PA      (GENO-) GENOME THERAPEUTICS CORP.
XX      XX
PI      Smith DR, Mao J;
XX      XX
DR      WPI; 2003-656441/62.
XX      XX
PT      New Mycobacterium tuberculosis anion pump peptide useful for as
PT      tuberculosis vaccine and diagnosis of tuberculosis infection.
XX      XX
PS      Disclosure; SEQ ID NO 131; 26pp; English.
XX      XX
CC      The invention relates to a non-naturally occurring peptide of
CC      Mycobacterium tuberculosis comprising an amino acid sequence
CC      corresponding to an anion pump protein. The invention also relates to a
CC      non-naturally occurring nucleic acid corresponding to a DNA sequence of
CC      Mycobacterium tuberculosis or Mycobacterium leprae. The new peptide is
CC      useful as a vaccine against Mycobacterium tuberculosis or Mycobacterium
CC      leprae or for screening for new tuberculosis drugs. Purified proteins

```

CC derived from the sequences of the invention may elicit a specific immune
CC response. The peptide may also be used to detect hypersensitivity
CC reactions of individuals exposed to Mycobacterium tuberculosis or
CC Mycobacterium leprae. The proteins and peptides may be affixed to solid
CC supports to detect antibodies typical of hypersensitivity reactions, from
CC a patient's sera. This sequence represents Mycobacterium leprae DNA of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification but was obtained in electronic format
CC directly from USPTO at seqdata.uspto.gov/sequence.html.
XX

SQ Sequence 42325 BP; 9673 A; 13128 C; 11330 G; 8194 T; 0 U; 0 Other;

Query Match 60.0%; Score 18; DB 10; Length 42325;
Best Local Similarity 80.8%; Pred. No. 5e+02;

Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 ACCCTCTGACTCGAGATTCCGCT 30
DB 688 ACCCACTGACTCCACAGTACGCT 713

RESULT 15

ADM27081_10/c

Continuation (11 of 17) of ADM27081 from base 1000001 (Hyperthermophile Methanopyrus kar
WP Sequence split into 17 fragments LOCUS ADM27081 Accession Adm27081

WP	Fragment Name	Begin	End
WP	ADM27081_00	1	110000
WP	ADM27081_01	100001	210000
WP	ADM27081_02	200001	310000
WP	ADM27081_03	300001	410000
WP	ADM27081_04	400001	510000
WP	ADM27081_05	500001	610000
WP	ADM27081_06	600001	710000
WP	ADM27081_07	700001	810000
WP	ADM27081_08	800001	910000
WP	ADM27081_09	900001	1010000
WP	ADM27081_10	1000001	1110000
WP	ADM27081_11	1100001	1210000
WP	ADM27081_12	1200001	1310000
WP	ADM27081_13	1300001	1410000
WP	ADM27081_14	1400001	1510000
WP	ADM27081_15	1500001	1610000
WP	ADM27081_16	1600001	1694968

Query Match 60.0%; Score 18; DB 11; Length 110000;
Best Local Similarity 80.8%; Pred. No. 5.4e+02;

Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CTGAGCCCTCTCGACTCGAGATTTC 26
DB 69684 CTTTCCGCTCTGACTCGAGATTTC 69659

Search completed: February 28, 2005, 23:51:44
Job time : 268.818 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2005, 23:32:36 ; Search time 875.065 Seconds
(without alignments)
1661.198 Million cell updates/sec

Title: US-10-086-062-4

Perfect score: 30
Sequence: 1 ctgagccctctcgactcgagagttccgct 30

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hgt: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sgs: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	30	6 AX358109	AX358109 Sequence
2	20.4	68.0	595	9 HUMQ76F01	AP085912 Homo sapi
3	20.4	68.0	98348	9 AL136103	AL136103 Human DNA
4	20.4	68.0	178965	9 AC012048	AC012048 Homo sapi
5	20.2	67.3	23211	3 AC099767	AC099767 Caenorhab
6	20	66.7	1503	1 AB038027	AB038027 Vibrio sp
7	20	66.7	156165	2 AC021453	AC021453 Homo sapi
8	20	66.7	157980	9 AC100852	AC100852 Homo sapi
9	19.6	65.3	25785	2 AC123011	AC123011 Rattus no
10	19.4	64.7	11406	4 AF107201	AF107201 Equus cab
11	19.4	64.7	70387	3 AC087075	AC087075 Caenorhab
12	19.4	64.7	110000	8 AC104791	AC104791 Homo sapi
13	19.4	64.7	159969	9 AC104791_6	Continuation (7 of
14	19.4	64.7	189269	2 AP006440	AP006440 Homo sapi
15	19.4	64.7	196472	2 AC011862	AC011862 Homo sapi
16	19.2	64.0	1500	1 AF118021	AF118021 Vibrio ca
17	19.2	64.0	1500	6 AR452360	AR452360 Sequence
18	19.2	64.0	1500	6 AX201179	AX201179 Sequence
19	63.3	93695	8	ATP2009	AL021749 Arabidops

c 20	19	63.3	156806	2	AC021353	AC021353 Homo sapi
c 21	19	63.3	177652	10	AC140251	AC140251 Mus muscu
c 22	19	63.3	199749	8	ATC81V68	AL161572 Arabidops
c 23	19	63.3	203120	10	AC124991	AC124991 Mus muscu
c 24	19	63.3	205475	2	AC130033	AC130033 Rattus no
c 25	19	63.3	213906	2	AC118634	AC118634 Mus muscu
c 26	19	63.3	223675	9	AC025031	AC025031 Homo sapi
c 27	19	63.3	232320	10	AC127314	AC127314 Mus muscu
c 28	19	63.3	244974	2	AC111249	AC111249 Rattus no
c 29	18.8	62.7	2481	6	AX827828	AX827828 Sequence
c 30	18.8	62.7	2481	10	RATR11BA2	W23366 Rattus norv
c 31	18.8	62.7	5550	1	AB016787	AB016787 Pseudomon
c 32	18.8	62.7	23107	6	AX695653	AX695653 Sequence
c 33	18.8	62.7	35758	2	AC127539	AC127539 Homo sapi
c 34	18.8	62.7	67145	2	AC124281	AC124281 Homo sapi
c 35	18.8	62.7	69833	2	AC010777	AC010777 Homo sapi
c 36	18.8	62.7	77691	9	AC120118	AC120118 Homo sapi
c 37	18.8	62.7	95241	9	HS39819	AL023096 Human DNA
c 38	18.8	62.7	100521	9	AC119397	AC119397 Homo sapi
c 39	18.8	62.7	101721	9	AC067950	AC067950 Homo sapi
c 40	18.8	62.7	110000	8	CR382131_15	Continuation (16 o
c 41	18.8	62.7	113196	9	HS3697K14	AL121829 Human DNA
c 42	18.8	62.7	141605	2	AC013732	AC013732 Homo sapi
c 43	18.8	62.7	146606	9	AC068213	AC068213 Homo sapi
c 44	18.8	62.7	148332	9	AC079586	AC079586 Homo sapi
c 45	18.8	62.7	148611	8	AC091680	AC091680 Oryza sat

ALIGNMENTS

RESULT 1	AX358109	Sequence 4 from Patent WO0194394.	30 bp	DNA	linear	PAT 13-FEB-2002
LOCUS	AX358109					
DEFINITION	AX358109					
ACCESSION	AX358109					
VERSION	AX358109.1	GI:18674856				
KEYWORDS						
SOURCE						
ORGANISM						
	Agrobacterium tumefaciens (Rhizobium radiobacter)					
	Agrobacterium tumefaciens					
	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;					
	Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.					
REFERENCE						
AUTHORS	1 Jilka, J.M., Hood, B.E. and Howard, J.A.					
TITLE	Novel plant promoter sequences and methods of use for same					
JOURNAL	Patent: WO 0194394-A 4 13-DEC-2001;					
	Prodigene, Inc. (US)					
FEATURES						
source	1..30					
	/organism="Agrobacterium tumefaciens"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:358"					
ORIGIN						
Query Match	100.0%; Score 30; DB 6; Length 30;					
Best Local Similarity	100.0%; Pred. No. 0.021;					
Matches	30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1 CTGAGACCCCTCTCGACTCGAGAGTTCCGCT 30					
Db	1 CTGAGACCCCTCTCGACTCGAGAGTTCCGCT 30					
RESULT 2						
HUMQ76F01/c	HUMQ76F01	595 bp	mRNA	linear		PRI 29-AUG-1998
LOCUS	HUMQ76F01					
DEFINITION	Homo sapiens full length insert cDNA clone YQ76F01.					
ACCESSION	AF085912					
VERSION	AF085912.1	GI:3483252				
KEYWORDS	FLI_CDNA.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					

REFERENCE 1 (bases 1 to 595)
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 Woensner,J., Tan,F., Marra,M., Kucaba,T., Yandell,M., Martin,J.,
 Marti,G., Bowles,L., Wylie,T., Bowers,Y., Steptoe,M., Theising,B.,
 Geisel,S., Allen,M., Underwood,K., Chappell,J., Person,B.,
 Gibbons,M., Harvey,N., Page,D., Chamberlain,A., Morales,R.,
 Schuck,R., Ritzer,E., Kohn,S., Swaller,T., Behmer,K., Hillier,L.,
 Wilson,R. and Waterston R.
 Full Clone Sequencing of the Longest Available Member from Each
 Unigene Cluster

TITLE
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 595)
 AUTHORS Waterston,R.
 TITLE Direct Submission
 JOURNAL Submitted (24-AUG-1998) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT SUBMITTED BY:
 Genome Sequencing Center
 Department of Genetics
 Washington University
 St. Louis MO 63108, USA
 http://genome.wustl.edu/gsc
 mailto:est@wustl.wustl.edu

NOTICE: This sequence represents the full insert of this cDNA. No attempt has been made to verify whether this corresponds to the full-length of the original mRNA from which it was derived. We have tried to obtain double-stranded, or double chemistry sequence across the entire clone, but potentially, there are areas in the sequence where this level of coverage was not achieved. Nevertheless, we are confident of the accuracy of this sequence as all regions of low quality, as defined by PHRAP (P. Green, in preparation), were visually inspected and edited accordingly. The consensus quality values for this sequence have been submitted separately.

FEATURES
 source The location of this clone is unknown.
 1..595 Location/Qualifiers

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:201721"
 /clone_1lb="Soares_fetal_liver_spleen_INFLS"
 9..338
 /rpt_family="L2"
 repeat_region

Query Match 68.0%; Score 20.4; DB 9; Length 595;
 Best Local Similarity 80.0%; Pred. No. 2.9e+02;
 Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 1 CTGAGCCCTCTGCTGAGCTCGGCT 30
 566 CTGAGCCCTCTGCTGAGCTCGGCT 537

RESULT 3
 LOCUS AL136103
 DEFINITION Human DNA sequence from clone RP1-250B11 on chromosome 10 contains
 STSs and GSSs, complete sequence.
 ACCESSION AL136103
 VERSION AL136103.24 GI:9581541
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 98348)
 AUTHORS Philimore,B.
 TITLE Direct Submission
 JOURNAL Submitted (04-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,

COMMENT

CB10 ISA, UK. E-mail enquiries: humquery@sanger.ac.uk
 Requests: clonerequest@sanger.ac.uk
 On Jul 26, 2000 this sequence version replaced gi:9501151.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPep; Information
 on the WormPep database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep
 This sequence is the entire insert of clone RP1-250B11. This sequence has been
 finished according to sequence map criteria as follows. An attempt
 is made to resolve all sequencing problems, such as compressions
 and repeats, but not necessarily within known annotated repeat
 sequence elements. Where the sequence is ambiguous, there is an
 annotation using the 'unsure' feature key. This sequence was
 generated from part of bacterial clone contigs of human chromosome
 10, constructed by the Sanger Centre Chromosome 10 Mapping Group.
 Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr10
 RP1-250B11 is from the library RP1-1 constructed by the group of
 Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pCYPAC2.

FEATURES
 source Location/Qualifiers

1..98348
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="10"
 /clone="RP1-250B11"
 /clone_1lb="RP1-1"
 1..34
 /note="Single clone region. Reads derived from clone PCR.
 Assembly consistent with restriction digest."
 2458..2765
 /note="AluX repeat: matches 1..308 of consensus"
 3087..3267
 /note="MER39b repeat: matches 355..546 of consensus"
 3269..3606
 /note="MER39 repeat: matches 13..380 of consensus"
 6106..6337
 /note="L2 repeat: matches 2435..2705 of consensus"
 6364..6662
 /note="AluX repeat: matches 1..298 of consensus"
 complement(8333..8799)
 /note="match: GSS: Em:AQ403443"
 complement(8375..8758)
 /note="match: GSS: Em:AQ264393"
 8802..9234
 /note="match: GSS: Em:AQ150662"
 9268..9356
 /note="match: STS: Em:G28019"
 11809..12092
 /note="AluX repeat: matches 1..284 of consensus"
 12205..12722
 /note="LIME repeat: matches 5285..5811 of consensus"
 14113..14171
 /note="L2 repeat: matches 2692..2750 of consensus"
 14548..14846
 /note="AluX repeat: matches 1..299 of consensus"
 15883..16169
 /note="AluX repeat: matches 1..287 of consensus"
 17146..18557
 /note="MER52A repeat: matches 1..1755 of consensus"
 18559..18713
 /note="MTR repeat: matches 6..160 of consensus"
 18955..19381
 /note="MER63 repeat: matches 590..1061 of consensus"

```
repeat_region 19384..19519
/note="FLAM C repeat: matches 3. .137 of consensus"
repeat_region 19713..19876
/note="MER63 repeat: matches 463. .577 of consensus"
repeat_region 19826..20117
/note="MER63 repeat: matches 2. .307 of consensus"
repeat_region 20325..20391
/note="MIR repeat: matches 88. .156 of consensus"
repeat_region 20440..20701
/note="MIR repeat: matches 2. .248 of consensus"
misc_feature 22321..22631
/note="match: STR: Em:HSJ10C4"
misc_feature 25725..26188
/note="match: GSS: Em:AQ211002"
misc_feature 25738..25978
/note="match: GSS: Em:AQ823705"
misc_feature 26548..27067
/note="match: GSS: Em:AQ798605"
misc_feature 26564..27250
/note="match: GSS: Em:AQ055423"
misc_feature 27549..27913
/note="match: GSS: Em:B87793"
repeat_region 28724..28952
/note="L2 repeat: matches 2167. .2416 of consensus"
misc_feature complement(29565..30037)
/note="match: GSS: Em:AQ155973"
misc_feature 30073..30670
/note="match: GSS: Em:AQ195052"
misc_feature complement(31752..32121)
/note="match: GSS: Em:AQ099305"
misc_feature complement(32344..32928)
/note="match: GSS: Em:AQ485233"
misc_feature 32930..33368
/note="match: GSS: Em:AQ424293"
repeat_region 34554..34583
/note="L5 copies 2 mer ca 100% conserved"
repeat_region 35059..35135
/note="L2 repeat: matches 2417. .2500 of consensus"
misc_feature complement(35681..36191)
/note="match: GSS: Em:B55923"
misc_feature complement(35777..36102)
/note="match: GSS: Em:AQ223744"
misc_feature complement(35861..36186)
/note="match: GSS: Em:AQ803876"
repeat_region 36692..36800
/note="MIR repeat: matches 13. .123 of consensus"
repeat_region 40461..40765
/note="AluX repeat: matches 1. .293 of consensus"
repeat_region 41312..41502
/note="MIR repeat: matches 8. .192 of consensus"
repeat_region 41785..41840
/note="28 copies 2 mer tg 85% conserved"
repeat_region 43022..43334
/note="AluX repeat: matches 3. .312 of consensus"
misc_feature complement(44477..45066)
/note="match: GSS: Em:AQ540223"
misc_feature complement(44671..45063)
/note="match: GSS: Em:AQ337658"
repeat_region 45077..45157
/note="MER57-internal repeat: matches 7151. .7230 of consensus"
repeat_region 45158..45304
/note="L1P repeat: matches 5050. .5197 of consensus"
repeat_region 45284..45621
/note="L1P4 repeat: matches 5476. .5812 of consensus"
misc_feature complement(45941..46452)
/note="match: GSS: Em:AQ832143"
repeat_region 48564..48709
/note="L2 repeat: matches 2596. .2749 of consensus"
repeat_region 48720..49082
/note="TH1B repeat: matches 1. .364 of consensus"
repeat_region 49176..49629
/note="MLTID repeat: matches 1. .502 of consensus"
```

```
repeat_region 49784..49912
/note="FLAM A repeat: matches 8. .132 of consensus"
repeat_region 49914..50164
/note="MIR repeat: matches 2. .261 of consensus"
repeat_region 52850..52918
/note="L2 repeat: matches 2635. .2706 of consensus"
misc_feature 53092..53613
/note="match: GSS: Em:AQ800606"
misc_feature 53110..53608
/note="match: GSS: Em:AQ172250"
misc_feature 53151..53604
/note="match: GSS: Em:AQ170181"
misc_feature 53165..53586
/note="match: GSS: Em:AQ545768"
misc_feature 53179..53355
/note="match: GSS: Em:AQ898990"
misc_feature 53378..53834
/note="match: GSS: Em:AQ898990"
repeat_region 53776..53958
/note="MIR repeat: matches 41. .250 of consensus"
misc_feature complement(53900..54343)
/note="match: GSS: Em:AQ400632"
repeat_region 53959..54203
/note="AluX repeat: matches 65. .308 of consensus"
repeat_region 54204..54216
/note="MIR repeat: matches 250. .260 of consensus"
misc_feature 54361..54767
/note="match: GSS: Em:AQ229256"
repeat_region 54372..54565
/note="MIR repeat: matches 13. .261 of consensus"
misc_feature 54440..54730
/note="match: GSS: Em:AQ084956"
misc_feature complement(55008..55535)
/note="match: GSS: Em:AQ556262"
misc_feature 55610..55935
/note="match: GSS: Em:AQ594863"
repeat_region 56871..57293
/note="L1MD repeat: matches 883. .1319 of consensus"
repeat_region 57297..57602
/note="AluO repeat: matches 1. .305 of consensus"
```

```
Query Match 68.0%; Score 20.4; DB 9; Length 98348;
Best Local Similarity 80.0%; Pred.No.2e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTGACCCCTCTCGACTCGAGGTTCCGCT 30
Db 62965 CTGAGCCCTCTCGGCTGAGACTTCCTCT 62994

RESULT 4
LOCUS AC012048 178965 bp DNA linear PRI 20-APR-2002
DEFINITION Homo sapiens chromosome 10 clone RP11-43N22, complete sequence.
ACCESSION AC012048
VERSION AC012048.11 GI:19744964
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 178965)
REFERENCE
AUTHORS Smith,D.R.
TITLE Genome Therapeutics Corporation Sequencing Center: Human Genome
Sequence Data
JOURNAL Unpublished
AUTHORS Smith,D.R.
TITLE Direct Submision
JOURNAL Submitted (19-OCT-1999) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
REFERENCE
AUTHORS Smith,D.R.
```

TITLE Direct Submission
 JOURNAL Submitted (16-AUG-2001) Genome Therapeutics Corporation, 100 Beaver
 Street, Waltham, MA 02453, USA
 REFERENCE 4 (bases 1 to 178965)
 AUTHORS Smith, D.R.
 JOURNAL Direct Submission
 Submitted (27-MAR-2002) Genome Therapeutics Corporation, 100 Beaver
 Street, Waltham, MA 02453, USA
 REFERENCE 5 (bases 1 to 178965)
 AUTHORS Smith, D.R.
 JOURNAL Direct Submission
 Submitted (28-MAR-2002) Genome Therapeutics Corporation, 100 Beaver
 Street, Waltham, MA 02453, USA
 REFERENCE 6 (bases 1 to 178965)
 AUTHORS Smith, D.R.
 JOURNAL Direct Submission
 Submitted (20-APR-2002) Genome Therapeutics Corporation, 100 Beaver
 Street, Waltham, MA 02453, USA
 COMMENT On Mar 27, 2002 this sequence version replaced gi:15193325.
 FEATURES
 source
 Location/Qualifiers
 1..178965
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="10"
 /clone="RP11-43N22"
 /clone_11b="RPCT-11"

ORIGIN
 Query Match 68.0%; Score 20.4; DB 9; Length 178965;
 Best Local Similarity 80.0%; Pred. No. 1.9e+02;
 Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTGAGACCCCTCTCGACTCGAGAGTTCCGCT 30
 Db 32913 CTGAGCCCTCTCGCTCGAGACTTCTCT 32942

RESULT 5
 AC099767/c 23211 bp DNA linear INV 20-NOV-2001
 LOCUS Caenorhabditis briggsae cosmid G24F01, complete sequence.
 ACCESSION AC099767
 VERSION AC099767.1 GI:17017639
 KEYWORDS HTG.
 SOURCE Caenorhabditis briggsae
 ORGANISM Caenorhabditis briggsae
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
 Rhabditioidea; Rhabditidae; Peloderinae; Caenorhabditis.
 REFERENCE 1 (bases 1 to 23211)
 AUTHORS Washington University Genome Sequencing Center.
 JOURNAL Direct Submission
 Submitted (20-NOV-2001) Department of Genetics, Washington
 University, Genome Sequencing Center, 444 Forest Park Avenue, St.
 Louis, MO 63110, USA
 REFERENCE 2 (bases 1 to 23211)
 AUTHORS Waterston, R.
 JOURNAL Direct Submission
 Submitted (20-NOV-2001) Department of Genetics, Washington
 University, Genome Sequencing Center, 444 Forest Park Avenue, St.
 Louis, MO 63110, USA
 COMMENT
 Genome Sequencing Center
 Department of Genetics, Washington University
 St. Louis, MO 63110, USA
 email: twmemetade.wustl.edu

NOTICE: This sequence may not be the entire insert of this clone.
 It may be shorter because we only sequence overlapping sections
 once, or longer because we provide a small overlap between
 neighboring submissions.

This sequence was finished as follows unless otherwise noted: all
 regions were double stranded, sequenced with an alternate chemistry
 or covered by high quality data (i.e., phred quality >= 30); an
 attempt was made to resolve all sequencing problems, such as
 compressions and repeats; all regions were covered by sequence from
 more than one m3 subclone.
 Location/Qualifiers
 1..23211
 /organism="Caenorhabditis briggsae"
 /mol_type="genomic DNA"
 /strain="Gujarat G16"
 /db_xref="taxon:6238"

ORIGIN
 Query Match 67.3%; Score 20.2; DB 3; Length 23211;
 Best Local Similarity 88.0%; Pred. No. 2.7e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 GACCCCTCTCGACTCGAGAGTTCCG 28
 Db 14462 GATTCCTCTCACTCGAGAGTTCCG 14438

RESULT 6
 AB038027/c 1503 bp DNA linear BCT 18-FEB-2000
 LOCUS Vibrio sp. SG128 gene for 16S rRNA, partial sequence.
 ACCESSION AB038027
 VERSION AB038027.1 GI:7007473
 KEYWORDS 16S ribosomal RNA.
 SOURCE Vibrio sp. SG128
 ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 Vibrionaceae; Vibrrio.
 REFERENCE 1 (bases 1 to 1503)
 AUTHORS Urakawa, H.
 JOURNAL 16S rRNA gene of marine bacterium
 Published Only in Database (2000)
 REFERENCE 2 (bases 1 to 1503)
 AUTHORS Urakawa, H.
 JOURNAL Direct Submission
 Submitted (05-FEB-2000) Hidetoshi Urakawa, Northwestern University,
 Department of Civil Engineering; Technological Institute 2145
 Sheridan Road, Evanston, Illinois 60208-3109, USA
 (E-mail: h-urakawa@nwu.edu, Tel: +1-847-467-5710,
 Fax: +1-847-491-4011)
 FEATURES
 source
 Location/Qualifiers
 1..1503
 /organism="Vibrio sp. SG128"
 /mol_type="genomic DNA"
 /strain="SG128"
 /db_xref="taxon:115126"
 <1..>1503
 /product="16S ribosomal RNA"

ORIGIN
 Query Match 66.7%; Score 20; DB 1; Length 1503;
 Best Local Similarity 82.1%; Pred. No. 4.1e+02;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CTGAGACCCCTCTCGACTCGAGAGTTCCG 28
 Db 231 CTGGGCCCATCCGACGCGAGAGTTCCG 204

RESULT 7
 AC021453/c 156165 bp DNA linear HTG 01-APR-2000
 LOCUS Homo sapiens clone RP11-125C16, WORKING DRAFT SEQUENCE, 14
 ACCESSION AC021453
 VERSION AC021453.3 GI:7382318

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens (human)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 156165)
 AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 TITLE Homo sapiens, clone RP11-125C16
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 156165)
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F., Boguski, L., Bouckgeater, B., Brown, A., Burkett, G., Castle, A., Chappel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Deaellano, K., Dewar, K., Domino, M., Doyle, M., Feneator, J., Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galegan, J., Garlyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, D., Landers, T., Lehocsky, J., Levine, R., Lien, C., Liu, G., Locke, K., MacDonald, P., Margulis, N., McEwan, P., McGurk, A., McKernan, K., McPheters, R., Meldrim, J., Menus, L., Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., Olivar, T.M., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rothman, D., Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.
 DIRECT SUBMISSION
 SUBMITTED (16-JAN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 ON APR 1, 2000 this sequence version replaced g1:6721267.
 ALL REPEATS WERE IDENTIFIED USING REPEATMASKER:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L5169
 Center clone name: 125_C16
 ----- Summary Statistics
 Sequencing vector: M13; M77815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 145142 bases at least Q40
 Consensus quality: 151009 bases at least Q30
 Consensus quality: 153345 bases at least Q20
 Insert size: 16000; agarose-fp
 Insert size: 154865; sum-of-contents
 Quality coverage: 4.1 in Q20 bases; agarose-fp
 Quality coverage: 4.2 in Q20 bases; sum-of-contents

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 14 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 1836: contig of 1836 bp in length
 * 1837 1936: gap of 100 bp
 * 1937 5796: contig of 3860 bp in length
 * 5797 5896: gap of 100 bp
 * 5897 9797: contig of 3901 bp in length
 * 9798 9897: gap of 100 bp
 * 9898 13919: contig of 4022 bp in length
 * 13920 14019: gap of 100 bp
 * 14020 17400: contig of 3381 bp in length
 * 17401 17500: gap of 100 bp
 * 17501 21253: contig of 3753 bp in length

FEATURES
 source
 1..156165
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="RP11-125C16"
 /clone_1ib="RP11-125C16 Human Male BAC"
 1..1836
 /note="assembly_fragment"
 1937..5796
 /note="assembly_fragment"
 5897..9797
 /note="assembly_fragment"
 9898..13919
 /note="assembly_fragment"
 14020..17400
 /note="assembly_fragment"
 17501..21253
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:right
 21354..31096
 /note="assembly_fragment"
 31197..41465
 /note="assembly_fragment"
 41566..50901
 /note="assembly_fragment"
 51002..62041
 /note="assembly_fragment"
 62142..75247
 /note="assembly_fragment"
 75348..89874
 /note="assembly_fragment"
 clone_end:T7
 vector_side:right
 89975..116362
 /note="assembly_fragment"
 116463..156165
 /note="assembly_fragment"
 21254 21353: gap of 100 bp
 * 21354 31096: contig of 973 bp in length
 * 31097 31197: gap of 100 bp
 * 31197 41465: contig of 10269 bp in length
 * 41466 41565: gap of 100 bp
 * 41566 50901: contig of 9336 bp in length
 * 50902 51001: gap of 100 bp
 * 51002 62041: contig of 11040 bp in length
 * 62042 62141: gap of 100 bp
 * 62142 75247: contig of 13106 bp in length
 * 75248 75347: gap of 100 bp
 * 75348 89874: contig of 14527 bp in length
 * 89875 89974: gap of 100 bp
 * 89975 116362: contig of 26388 bp in length
 * 116363 116462: gap of 100 bp
 * 116463 156165: contig of 39703 bp in length.
 Location/Qualifiers
 1..156165
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="RP11-125C16"
 /clone_1ib="RP11-125C16 Human Male BAC"
 1..1836
 /note="assembly_fragment"
 1937..5796
 /note="assembly_fragment"
 5897..9797
 /note="assembly_fragment"
 9898..13919
 /note="assembly_fragment"
 14020..17400
 /note="assembly_fragment"
 17501..21253
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:right
 21354..31096
 /note="assembly_fragment"
 31197..41465
 /note="assembly_fragment"
 41566..50901
 /note="assembly_fragment"
 51002..62041
 /note="assembly_fragment"
 62142..75247
 /note="assembly_fragment"
 75348..89874
 /note="assembly_fragment"
 clone_end:T7
 vector_side:right
 89975..116362
 /note="assembly_fragment"
 116463..156165
 /note="assembly_fragment"
 ORIGIN
 Query Match 66.7%; Score 20; DB 2; Length 156165;
 Best Local Similarity 82.1%; Pred. No. 2.9e+02;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Oy 2 TGGACCTCTCGACTCGAGAGTTCGC 29
 Db 11285 TGGACCTCTCTATTCAAGAGTTCGC 11258
 RESULT 8
 AC100852/c 157980 bp DNA linear PRI 29-AUG-2002
 LOCUS Homo sapiens chromosome 17, clone RP11-125C16, complete sequence.
 DEFINITION AC100852
 ACCESSION AC100852.2 GI:22539166
 VERSION
 KEYWORDS HTG.

```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 157980)
AUTHORS     Birren,B., Nussbaum,C. and Lander,E.
JOURNAL     Homo sapiens chromosome 17, clone RP11-125C16
TITLE       Unpublished
REFERENCE   2 (bases 1 to 157980)
AUTHORS     Birren,B., Linton,L., Nussbaum,C., Lander,E., All,A., Allen,N.,
JOURNAL     Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B.,
TITLE       Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
AUTHORS     Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
JOURNAL     Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
TITLE       Ferrerita,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
AUTHORS     Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
JOURNAL     Hages,B., Heaford,A., Horton,L., Hulme,W., Illiv,I., Johnson,R.,
TITLE       Jones,C., Kamat,A., Karatas,A., Kelle,C., Lacroque,K.,
AUTHORS     Lamazeres,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
JOURNAL     Maclean,C., Macdonald,P., Major,J., Margolis,N., Meldrum,J.,
TITLE       McCarthy,M., McEwan,P., McKernan,K., McPheters,R., Meldrum,J.,
AUTHORS     Menus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
JOURNAL     Norzu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
TITLE       Oliver,J., Peterson,K., Phukhang,P., Pierre,N., Pollara,V.,
AUTHORS     Raymond,C., Retta,R., Rieback,W., Riley,R., Rise,C., Rogov,P.,
JOURNAL     Roman,J., Roetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
TITLE       Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
AUTHORS     Strauss,N., Subramanian,A., Talamas,J., Teefaye,S., Theodore,J.,
JOURNAL     Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
TITLE       Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
AUTHORS     Zainoun,J., Zemke,L., Zimmer,A. and Zody,M.
JOURNAL     Direct Submission
TITLE       Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
AUTHORS     Research, 320 Charles Street, Cambridge, MA 02141, USA
JOURNAL     3 (bases 1 to 157980)
REFERENCE   3 (bases 1 to 157980)
AUTHORS     Birren,B., Nussbaum,C., Lander,E., All,A., Allen,N., Anderson,S.,
JOURNAL     Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,
TITLE       Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
AUTHORS     Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
JOURNAL     Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hages,B.,
TITLE       Horton,L., Hulme,W., Illiv,I., Johnson,R., Jones,C., Kamat,A.,
AUTHORS     Karatas,A., Kelle,C., Landers,T., Levine,R., Lindblad-Toh,K.,
JOURNAL     Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
TITLE       McCarthy,M., Meldrum,J., Menus,L., Mihova,T., Mlenga,V.,
AUTHORS     Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
JOURNAL     O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
TITLE       Phukhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
AUTHORS     Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P.,
JOURNAL     Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
TITLE       Teefaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
AUTHORS     Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
JOURNAL     Zemke,L., Zimmer,A. and Zody,M.
JOURNAL     Direct Submission
TITLE       Submitted (29-AUG-2002) Whitehead Institute/MIT Center for Genome
AUTHORS     Research, 320 Charles Street, Cambridge, MA 02141, USA
JOURNAL     On Aug 29, 2002 this sequence version replaced gi:17046222.
COMMENT     All repeats were identified using RepeatMasker:
COMMENT     Smit, A.F.A. & Green, P. (1996-1997)
COMMENT     http://ftp.genome.washington.edu/RM/RepeatMasker.html
JOURNAL     ----- Genome Center
JOURNAL     Center: Whitehead Institute/ MIT Center for Genome Research
JOURNAL     Center code: WtBR
JOURNAL     Web site: http://www-seq.wi.mit.edu
JOURNAL     Contact: sequence_submissions@genome.wi.mit.edu
JOURNAL     ----- Project Information
JOURNAL     Center project name: L21587
JOURNAL     Center clone name: 125_C_16
JOURNAL     -----
JOURNAL     Location/Qualifiers
JOURNAL     1..157980
JOURNAL     /organism="Homo sapiens"
JOURNAL     /mol_type="genomic DNA"

```

```

/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-125C16"
/clone_11b="RP11-125C16 Human Male BAC"
62..173
/rpt_family="MIR"
repeat_region
complement(636..832)
/rpt_family="MIR"
repeat_region
833..888
/rpt_family="MER3"
complement(891..1207)
/rpt_family="AluDb"
repeat_region
1208..1283
/rpt_family="MER3"
1232..1236
/nc="30 qual SINGL region"
complement(1303..2370)
/rpt_family="L1PA4"
1330..1354
/nc="30 qual SINGL region"
2371..3213
/rpt_family="L1PA4"
complement(3266..3408)
/rpt_family="Charlie7"
3506..3538
/rpt_family="L2"
3539..3653
/rpt_family="MER47A"
complement(3654..3926)
/rpt_family="AluSc"
3927..4141
/rpt_family="MER47A"
4142..4438
/rpt_family="AluY"
4439..4473
/rpt_family="MER47A"
4474..4561
/rpt_family="L2"
complement(4601..4700)
/rpt_family="L2"
complement(4708..4845)
/rpt_family="MER5A"
complement(4944..5104)
/rpt_family="MIR"
complement(5112..5532)
/rpt_family="L2"
complement(5546..5822)
/rpt_family="AluDb"
complement(5823..5999)
/rpt_family="L2"
6000..6216
/rpt_family="AluDb"
complement(6217..6292)
/rpt_family="L2"
6447..6752
/rpt_family="AluDb"
7412..7472
/rpt_family="A-rich"
7788..7907
/rpt_family="CGG"
8642..8683
/rpt_family="T-rich"
8954..8978
/rpt_family="CCA"
9084..9123
/rpt_family="CAG"
10990..11017
/rpt_family="AT-rich"
11225..11326
complement(11372..11504)
/rpt_family="MIR1B"

```


genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GYMD
Center clone name: CH230-92124

----- Summary Statistics

Assembly program: Phrap; version 0.990329
Consensus quality: 236342 bases at least Q40
Consensus quality: 239870 bases at least Q30
Consensus quality: 242005 bases at least Q20
Estimated insert size: 243903; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 250677: contig of 250677 bp in length
* 250678 250777: gap of unknown length
* 250778 252307: contig of 1530 bp in length
* 252308 252407: gap of unknown length
* 252408 253577: contig of 1170 bp in length
* 253578 253678: gap of unknown length
* 253678 257595: contig of 3918 bp in length.

FEATURES

source

1..257595

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-92124"

/complement(246343..247143)

/note="clone boundary"

clone end:Sp6

site:EcORI

end_sequence: BH295371"

misc_feature

end_sequence: BH295371"

complement(246398..247079)

/note="clone boundary"

clone end:Sp6

site:EcORI

end_sequence: BH295371"

ORIGIN

Query Match 65.3%; Score 19.6; DB 2; Length 257595;
Best Local Similarity 84.6%; Pred. No. 4.2e+02;

Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTGGACCCCTCTGACTCGAGAGTTCC 26
Db 151391 CTGGACCTCTTTCAGCCCTAGAGTTCC 151416

RESULT 10
AF107201/c 11406 bp DNA linear MAM 13-DEC-1998

DEFINITION Equus caballus beta-lactoglobulin II (BLG) gene, complete cds.
ACCESSION AF107201
VERSION AF107201.1 GI:4008111

KEYWORDS Equus caballus (horse)

SOURCE Equus caballus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

1 (bases 1 to 11406)

AUTHORS Masei, A.M., Brandon, R.B. and Bell, T.K.

JOURNAL Nucleotide Sequence of the Equine Beta-Lactoglobulin Gene

REFERENCE Unpublished

AUTHORS 2 (bases 1 to 11406)

Masei, A.M., Brandon, R.B. and Bell, T.K.

JOURNAL Direct Submission

Submitted (17-NOV-1998) Australian Equine Blood Typing Research

Laboratory, University of Queensland, St. Lucia, Brisbane,

Queensland 4072, Australia

FEATURES

source

1..11406

/organism="Equus caballus"

/mol_type="genomic DNA"

/db_xref="taxon:9796"

<4472..>8679

/gene="BLG"

join(4472..4567,4956..5095,5767..5840,7367..7480,

8302..8406,8663..8679)

/gene="BLG"

/product="beta-lactoglobulin II"

join(4472..4567,4956..5095,5767..5840,7367..7480,

8302..8406,8663..8679)

/gene="BLG"

/codon_start=1

/product="beta-lactoglobulin II"

/protein_id="AAC95385.1"

/db_xref="GI:4008112"

/translation="MKCLILALGSLMGNOATDIPOTMODLLOEVAGRWHSYANMA

SPISLDSRSVPLRVYVEIRPTPGNFIIRBANNAICERINVAQTEDPVFTV

NYGGRKTSVLDTPYAHYFRCVGPPLPSAEHGVCOYLARTQVDEWKEFRALQ

PIPRGVQIVQDPSGGQERCG"

ORIGIN

Query Match 64.7%; Score 19.4; DB 4; Length 11406;

Best Local Similarity 79.3%; Pred. No. 6.5e+02;

Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CTGGACCCCTCTGACTCGAGAGTTCCG 29
Db 6919 CGGGGCCCTCTTGAGTCAGAGTTACAC 6891

RESULT 11

LOCUS

AC087075 70387 bp DNA linear INV 05-DEC-2000

DEFINITION Caenorhabditis briggsae cosmid CB023K10, complete sequence.

AC087075 AC087075.1 GI:11545986

VERSION HTG.

KEYWORDS Caenorhabditis briggsae

SOURCE Caenorhabditis briggsae

ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;

Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;

Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.

1 (bases 1 to 70387)

AUTHORS Washington University Genome Sequencing Center.

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 70387)

Waterston, R.

Direct Submission

Submitted (05-DEC-2000) Department of Genetics, Washington

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

Submitted by:

Genome Sequencing Center

Department of Genetics, Washington University,

St. Louis, MO 63110, USA

e-mail: jspleth@watson.wustl.edu

COMMENT

NOTICE: This sequence may not be the entire insert of this clone.
It may be shorter because we only sequence overlapping sections

once, or longer because we provide a small overlap between neighboring submissions.

FEATURES

source

Location/Qualifiers

1. 70387
/organism="Caenorhabditis briggsae"
/mol_type="genomic DNA"
/strain="Gujarat G16"
/db_xref="taxon:6238"
/clone="CB023K10"
23139. 23209
/product="tRNA-Gly"
/note="codon recognized: GGC"
complement(36569. 36640)
/product="tRNA-Glu"
/note="codon recognized: GAG"
37437. 37508
/product="tRNA-Gln"
/note="codon recognized: CAA"

ORIGIN

Query Match 64.7%; Score 19.4; DB 3; Length 70387;
Best Local Similarity 79.3%; Pred. No. 5.7e+02;

Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY

2 TGGACCCCTCGACTCGAGATTCCGCT 30

Db 6252 TGGCAGCTCTCGATTGAAAGTTGAGCT 6280

RESULT 12

AE016816 6/c

WPCOMMENT

Sequence split into 9 fragments LOCUS AE016816 Accession AE016816

Fragment Name	Begin	End
AE016816_0	1	110000
AE016816_1	100001	210000
AE016816_2	200001	310000
AE016816_3	300001	410000
AE016816_4	400001	510000
AE016816_5	500001	610000
AE016816_6	600001	710000
AE016816_7	700001	810000
AE016816_8	800001	907057

Continuation (7 of 9) of AE016816 from base 600001 (AE016816 Eremothecium goesypii ATCC

Query Match 64.7%; Score 19.4; DB 8; Length 110000;

Best Local Similarity 79.3%; Pred. No. 5.5e+02;

Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTGGACCCCTCGACTCGAGATTCCGC 29

Db 20976 CTGGACTGCTGTCGATTCGAGACTTCAC 20948

RESULT 13

AC104791/c 159969 bp DNA linear PRI 21-FEB-2002

LOCUS AC104791 Homo sapiens BAC clone RP11-181K12 from 4, complete sequence.

DEFINITION AC104791 AC032008

AC104791.3 GI:18482313

HTG.

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 159969)

1 Sulston, J.E. and Watson, R.

Toward a complete human genome sequence

Genome Res. 8 (11), 1097-1108 (1998)

JOURNAL

GENOME

RESEARCH

2 (bases 1 to 159969)

1sak, A., Meyer, R. and Creason, K.

TITLE The sequence of Homo sapiens BAC clone RP11-181K12

JOURNAL

Unpublished (2001)

REFERENCE

AUTHORS

TITLE

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

```

/chromosome="4"
/map="4"
/clone="RP11-181K12"
/clone_lib="RPCT-11"
191. .463
/rpt_family="MIR"
repeat_region
1855. .2217
/rpt_family="MALR"
repeat_region
2333. .2362
/rpt_family="AT_rich"
repeat_region
2471. .2690
/rpt_family="MIR"
repeat_region
3143. .3169
/rpt_family="(A)n"
repeat_region
3296. .3316
/rpt_family="AT_rich"
repeat_region
3304. .3462
/rpt_family="Alu"
repeat_region
3589. .3609
/rpt_family="AT_rich"
repeat_region
3984. .4033
/rpt_family="ERV_L"
repeat_region
4019. .4111
/rpt_family="(TA)n"
repeat_region
4087. .4495
/rpt_family="ERV_L"
repeat_region
5195. .5483
/rpt_family="Alu"
repeat_region
5374. .5402
/rpt_family="(GGATG)n"
repeat_region
5897. .5927
/rpt_family="(TTTTG)n"
repeat_region
5906. .6217
/rpt_family="Alu"
repeat_region
6577. .6760
/nc="similar to Homo sapiens EST AW843469
(NID:g7937452)"
7129. .7131
/nc="similar to Homo sapiens EST AW843469
(NID:g7937452)"
7314. .7393
/rpt_family="MIR"
repeat_region
7394. .7415
/rpt_family="(T)n"
repeat_region
7700. .7766
/rpt_family="MIR"
repeat_region
9154. .9438
/rpt_family="Alu"
repeat_region
9437. .9466
/rpt_family="(TAA)n"
repeat_region
10074. .10204
/rpt_family="MIR"
repeat_region
10351. .10812
/rpt_family="MALR"
repeat_region
11128. .11343
/rpt_family="Alu"
repeat_region
11414. .11626
/rpt_family="MIR_type"
repeat_region
13354. .13449
/rpt_family="MIR_type"
repeat_region
13450. .13721
/rpt_family="Alu"
repeat_region
13722. .13784
/rpt_family="MIR_type"
repeat_region
13813. .13852
/rpt_family="CT-rich"
repeat_region
15896. .15941
/rpt_family="(TG)n"
repeat_region
16404. .16916
/rpt_family="L2"
repeat_region
16918. .16939
/rpt_family="(T)n"
repeat_region
18098. .18220

```

```

/rpt_family="L1"
18250. .18539
/rpt_family="Alu"
repeat_region
18545. .18594
/rpt_family="(A)n"
misc_feature
18829. .19205
/nc="match to EST AA743144 (NID:g2782650) my20h09.s1"
repeat_region
19565. .19869
/rpt_family="Alu"
repeat_region
20046. .20785
/rpt_family="ERV_L"
repeat_region
21452. .21689
/rpt_family="L1"
repeat_region
21764. .22061
/rpt_family="Alu"
repeat_region
21780. .22172
/nc="match to EST AA554485 (NID:g2325024) n136b03.s1"
repeat_region
22691. .22990
/rpt_family="Alu"
repeat_region
23268. .23312
/rpt_family="(TTTA)n"
repeat_region
23286. .23596
/rpt_family="Alu"
repeat_region
23637. .24312
/rpt_family="MIR2_type"
repeat_region
24463. .24756
/rpt_family="L2"
repeat_region
24798. .25023
/rpt_family="MIR2_type"
repeat_region
25050. .25411
/rpt_family="ERV_L"
repeat_region
25377. .25406
/rpt_family="(CAA)n"
repeat_region
25412. .25733
/rpt_family="Alu"
repeat_region
25532. .25561
/rpt_family="(A)n"
repeat_region

```

Query Match 64.7%; Score 19.4; DB 9; Length 159969;
 Best Local Similarity 79.3%; Pred. No. 5.3e+02;
 Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 TGGACCCCTCGACTCGAGATTCCGCT 30
 Db 120906 TGGACTACTCTGCTCCGACTTCTCT 120878

RESULT 14
 AP006440/c 189269 bp DNA linear HTG 21-MAY-2003
 LOCUS Homo sapiens chromosome 11 clone RP11-368120 map 11q, WORKING DRAFT
 DEFINITION SEQUENCE, 4 unordered pieces.
 ACCESSION AP006440
 VERSION AP006440.1 GI:30962586
 KEYWORDS HTG, HTGS_PHASE1, HTGS_DRAFT.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 TITLE Homo sapiens genomic DNA of 11q
 JOURNAL Published Only in Database (2003)
 REFERENCE 2 (bases 1 to 189269)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (19-MAY-2003) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail:hattori@gsc.riken.go.jp URL:http://hgp.gsc.riken.go.jp/,
 Tel:81-45-503-9111, Fax:81-45-503-9170)

COMMENT

----- Genome Center
Center: RIKEN Genomic Sciences Center (GSC)

Center code: RIKEN

Web site: <http://hgp.gsc.riken.go.jp/>
Contact: hatori@gsc.riken.go.jp

----- Project Information
Center project name: HumDrafc11
Center clone name: RP11-368120

----- Summary Statistics

Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329

Consensus quality: 188057 bases at least Q40

Consensus quality: 188482 bases at least Q30

Consensus quality: 188825 bases at least Q20

Insert size: 188969; sum-of-contigs

Quality coverage: 12.20x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1 70266 contig of 70266 bp in length
70367 130873 contig of 60507 bp in length
130974 188629 contig of 57656 bp in length
188730 189269 contig of 540 bp in length.
* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
* 1 70266 contig of 70266 bp in length
* 70367 130873 contig of 60507 bp in length
* 130974 188629 contig of 57656 bp in length
* 188730 189269 contig of 540 bp in length.
* Location/Qualifiers

FEATURES
source
1. 189269
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-368120"

misc_feature
1. 70266
/note="assembly_fragment"

misc_feature
70367..130873
/note="assembly_fragment"

misc_feature
130974..188629
/note="assembly_fragment"

misc_feature
188730..189269
/note="assembly_fragment"

ORIGIN

Query Match 64.7%; Score 19.4; DB 2; Length 189269;
Best Local Similarity 79.3%; Pred. No. 5.3e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 TGAGACCCCTCGACTCGAGAGTTCGCT 30
|||||
Db 161287 TGAGACTACTCTCTCTCCAGACTTCTCT 161259

RESULT 15
AC011862/c

LOCUS

AC011862 196472 bp DNA linear HTG 13-JUL-2000

DEFINITION
AC011862 Homo sapiens clone RP11-15K8, LOW-PASS SEQUENCE SAMPLING.

VERSION
AC011862.4 GI:7144913

KEYWORDS
HTG; HTGS_PHASE0.

SOURCE
Homo sapiens (human)

ORGANISM
Homo sapiens

REFERENCE
1 (bases 1 to 196472)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.

REFERENCE
2 (bases 1 to 196472)
Unpublished
Homo sapiens chromosome, clone RP11-15K8

TITLE
JOURNAL
COMMENT

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baldwin, J., Barn, N., Beckerly, R., Boguslavsky, L., Boukhgalter, B., Brown, A., Casile, A., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearrellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M., Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J.C., Johnson, R., Jones, C., Kam, L., Karas, A., Klein, J., Lehoczy, J., Lien, C., Locke, K., MacDonald, P., Margulis, N., McEwan, P., McGuirk, A., McKernan, K., McLaughlin, D., Meldrum, J., Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (15-OCT-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 3, 2000 this sequence version replaced g1:6980935.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RW/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: L3484
Center clone name: 15_K_8

* NOTE: This record contains 217 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 883: contig of 883 bp in length
884 983: gap of 100 bp
884 1877: contig of 894 bp in length
1878 1977: gap of 100 bp
1978 2866: contig of 889 bp in length
2867 2966: gap of 100 bp
2967 3829: contig of 863 bp in length
3830 3929: gap of 100 bp
3930 4831: contig of 902 bp in length
4832 4931: gap of 100 bp
4932 5790: contig of 859 bp in length
5791 6752: contig of 862 bp in length
6753 6853: gap of 100 bp
6854 7749: contig of 897 bp in length
7750 7849: gap of 100 bp
7850 8754: contig of 905 bp in length
8755 8854: gap of 100 bp

```
* 8855 9749: contig of 895 bp in length
* 9750 9849: gap of 100 bp
* 9850 10738: contig of 889 bp in length
* 10739 10838: gap of 100 bp
* 10839 11725: contig of 887 bp in length
* 11726 11825: gap of 100 bp
* 11826 12707: contig of 882 bp in length
* 12708 12807: gap of 100 bp
* 12808 13697: contig of 890 bp in length
* 13698 13797: gap of 100 bp
* 13798 14668: contig of 871 bp in length
* 14669 14768: gap of 100 bp
* 14769 15648: contig of 880 bp in length
* 15649 15748: gap of 100 bp
* 15749 16622: contig of 874 bp in length
* 16623 16722: gap of 100 bp
* 16723 17587: contig of 865 bp in length
* 17588 17687: gap of 100 bp
* 17688 18558: contig of 871 bp in length
* 18559 18658: gap of 100 bp
* 18659 19534: contig of 876 bp in length
* 19535 19634: gap of 100 bp
* 19635 20510: contig of 876 bp in length
* 20511 20610: gap of 100 bp
* 20611 21517: contig of 907 bp in length
* 21518 21617: gap of 100 bp
* 21618 22499: contig of 882 bp in length
* 22500 22599: gap of 100 bp
* 22600 23480: contig of 881 bp in length
* 23481 23580: gap of 100 bp
* 23581 24444: contig of 864 bp in length
* 24445 24544: gap of 100 bp
* 24545 25479: contig of 935 bp in length
* 25480 25580: gap of 100 bp
* 25580 26492: contig of 913 bp in length
* 26493 26592: gap of 100 bp
* 26593 27454: contig of 862 bp in length
* 27455 27554: gap of 100 bp
* 27555 28431: contig of 877 bp in length
* 28432 28531: gap of 100 bp
* 28532 29495: contig of 964 bp in length
* 29496 29595: gap of 100 bp
* 29596 30489: contig of 894 bp in length
* 30490 30589: gap of 100 bp
* 30590 31436: contig of 846 bp in length
* 31437 31535: gap of 100 bp
* 31536 32425: contig of 890 bp in length
* 32426 32525: gap of 100 bp
* 32526 33401: contig of 876 bp in length
* 33402 33501: gap of 100 bp
* 33502 34376: contig of 875 bp in length
* 34377 34476: gap of 100 bp
* 34477 35355: contig of 879 bp in length
* 35356 35455: gap of 100 bp
* 35456 36370: contig of 915 bp in length
* 36371 36470: gap of 100 bp
* 36471 37358: contig of 888 bp in length
* 37359 37458: gap of 100 bp
* 37460 38332: contig of 874 bp in length
* 38333 38432: gap of 100 bp
* 38433 39285: contig of 853 bp in length
* 39286 39385: gap of 100 bp
* 39386 40317: contig of 932 bp in length
* 40318 40417: gap of 100 bp
* 40418 41314: contig of 897 bp in length
* 41315 41414: gap of 100 bp
* 41415 42297: contig of 883 bp in length
* 42298 42397: gap of 100 bp
* 42398 43270: contig of 873 bp in length
* 43271 43370: gap of 100 bp
* 43371 44268: contig of 898 bp in length
* 44269 44368: gap of 100 bp
* 44370 45233: contig of 865 bp in length
```

```
* 45234 45333: gap of 100 bp
* 45334 46202: contig of 869 bp in length
* 46203 46302: gap of 100 bp
* 46303 47193: contig of 891 bp in length
* 47194 47293: gap of 100 bp
* 47294 48183: contig of 885 bp in length
* 48183 48282: gap of 100 bp
* 48283 49146: contig of 864 bp in length
* 49147 49246: gap of 100 bp
* 49247 50116: contig of 870 bp in length
* 50117 50216: gap of 100 bp
* 50217 51112: contig of 896 bp in length
* 51113 51212: gap of 100 bp
* 51213 52102: contig of 890 bp in length
* 52103 52202: gap of 100 bp
* 52203 53064: contig of 862 bp in length
* 53065 53164: gap of 100 bp
* 53165 54024: contig of 860 bp in length
* 54025 54124: gap of 100 bp
* 54125 54957: contig of 833 bp in length
* 54958 55057: gap of 100 bp
* 55058 55938: contig of 881 bp in length
* 55939 56038: gap of 100 bp
* 56039 56932: contig of 894 bp in length
* 56933 57032: gap of 100 bp
* 57033 57916: contig of 884 bp in length
* 57917 58016: gap of 100 bp
* 58017 58875: contig of 859 bp in length
* 58876 58975: gap of 100 bp
* 58976 59844: contig of 869 bp in length
* 59845 59944: gap of 100 bp
* 59945 60795: contig of 851 bp in length
* 60796 60895: gap of 100 bp
* 60896 61774: contig of 879 bp in length
* 61775 61875: gap of 100 bp
* 61876 62739: contig of 865 bp in length
* 62740 62839: gap of 100 bp
* 62840 63807: contig of 968 bp in length
* 63808 63907: gap of 100 bp
* 63908 64783: contig of 876 bp in length
* 64784 64883: gap of 100 bp
* 64884 65767: contig of 884 bp in length
* 65768 65867: gap of 100 bp
* 65868 66765: contig of 898 bp in length
* 66766 66865: gap of 100 bp
* 66866 67730: contig of 865 bp in length
* 67731 67830: gap of 100 bp
* 67831 68721: contig of 891 bp in length
* 68722 68821: gap of 100 bp
* 68822 69711: contig of 890 bp in length
* 69712 69811: gap of 100 bp
* 69812 70699: contig of 888 bp in length
* 70700 70799: gap of 100 bp
* 71679: contig of 880 bp in length
```

```
Query Match 64.7% Score 19.4; DB 2; Length 196472;
Best Local Similarity 79.3%; Pred. No. 5.3e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CTGACCCCTCTGACCTCGAGAGTTCCGC 29
Db 6946 CTGACCCCTGATGCTGAGAGTTCCCC 6918
```

Search completed: March 1, 2005, 00:29:18
Job time : 883.065 secs

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☒ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

This Page Blank (uspto)